THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume 38 January, 1947 Number 1

TIME RELATIONS FOR EYE-MOVEMENT MEASURES IN READING*

MILES A. TINKER
University of Minnesota

Two important aspects of oculomotor behavior in reading are fixational pauses and the interfixation or saccadic eye movements. Reading time consists of the time taken for the fixational pauses plus that for the saccadic movements. The pauses are relatively long in duration in comparison with movement time. It is well known that reading perception occurs only during the pauses. The interfixation movements are so rapid that clear vision is precluded.

These saccadic movements are employed not only in reading, but also in shifting the line of regard from one object to another in the field of view as in viewing a picture, a landscape or in inspecting factory products. Fundamentally, saccadic movements are reactions to eccentric stimulation (Dodge,[2]). There is a strong tendency to move the eyes so that the new point of interest will be seen with the foveas. As pointed out by Miles,[8] normally one cannot control the rate of saccadic movements. In shifting the eye from one fixation to another, the initiation of the movement is a voluntary affair. But after deciding to shift, the rate of the shift or the speed of the movement that occurs is apparently not under voluntary control. An analysis of horizontal saccadic movements by Dodge[2] revealed three distinct phases: first there is a positive acceleration to the maximum velocity; this maximum velocity is maintained for a considerable angle of movement and constitutes the second phase; this second phase gives way to the third phase, a negative

* The writer is grateful to the University of Minnesota Graduate School for research grant to finance this study.

acceleration as the eye comes to rest. For relatively short excursions of the eye, the second phase (period of maximum velocity) is very short. But in the relatively long excursions, the second phase is highly conspicuous. Confirmation of this analysis is found in the report by Miles.[8] Evidence presented by Dodge and Cline[3] and by Miles[8] shows that abductive (toward temple) saccadic movements are faster than adductive (toward nose) movements. For any individual eye, and for movement in the same direction (abduction or adduction) through a given angle, the duration of successive movements remains remarkably constant (Dodge,[2]).

As pointed out by Tinker,[11] the early determinations of speed of saccadic eye movements by Lamansky,[6] Erdmann and Dodge,[4] and Huey[5] were inaccurate because of faulty methodology. Only after the invention of the corneal-reflection method of recording eye movements, described by Dodge and Cline,[3] became available, was it possible to make accurate measurements of speed of eye movements. Inasmuch as perception in reading occurs during the fixational pause time, pauses rather than eye movements have received most of the attention in investigations of reading. Erdmann and Dodge,[4] however, in discussing the relation between movement duration and pause duration, pointed out that a large portion of reading time is devoted to the pauses. During the past fifteen years, the author, working at the University of Minnesota, has conducted several studies dealing with eye movement time and pause duration. The Minnesota eye movement camera[10] was employed to record responses. This apparatus is an adaptation of the Dodge corneal-reflection technique. The purpose of the present paper is to coördinate the results reported in these investigations in order to bring out the conditions determining the time taken for eye movements and the relation of this time to duration of pauses in reading.

Most of the data on the velocity of saccadic eye movements have been obtained from photographing eye movements which are made in response to peripheral stimuli as the eye looks back and forth between two fixation points. The character of such movements, however, is not materially different, qualitatively or quantitatively, from eye movements in reading according to

data presented by Dodge and Cline.[3] The results of these studies, therefore, are directly applicable to discussions of oculomotor behavior in reading.

Early studies on eye movement duration, pause duration and reading time have been summarized and evaluated by Tinker.[11] His report revealed that the information concerning angle velocity of saccadic eye movements was meager. For the most part, data on this subject have been obtained incidently in such studies as those of Benedict, Miles, Roth and Smith,[1] Miles,[7] and Travis.[14] The only systematic study of saccadic eye movement speed as such, reported by Dodge and Cline,[3] employed only a few measurements on their three subjects and made no attempt to evaluate individual differences or the stability of the obtained means.

More recently, Tinker[13] completed a systematic study of speed of horizontal saccadic eye movements. Fourteen college students, seven men and seven women, served as subjects. Data were obtained for horizontal excursion through five, ten, fifteen, twenty, twenty-five, thirty and thirty-five degrees of visual angle. There were ten records for each subject for each angle. A Latin square experimental design was employed and analysis of variance applied to the data. For any single individual there was slight variation from measurement to measurement for any particular angle of movement. No significant sex differences in speed of the saccadic movements occurred. Differences between individuals, while not large, were statistically significant.

Mean eye-movement speeds for the different angles, together with the data reported by Dodge and Cline,[3] are given in Table 1. Speeds are in milliseconds. The agreement of these data with the 1901 results of Dodge and Cline is remarkably close in view of the fact that individual differences are significant and that the earlier study used only three subjects. Other miscellaneous data fall in line with those obtained in this study. For visual angles of about 14 to 16 degrees, Tinker[11] obtained an angular velocity of 48 milliseconds. Shen[9] cites a speed of 54 milliseconds for an angle of about twenty degrees. In the same study, Shen[9] derived from Dodge and Cline's[3] data the estimated speeds for small angles:

9.9 milliseconds for 1 degree
17.3 milliseconds for 2 degrees
22.2 milliseconds for 3 degrees
26.0 milliseconds for 4 degrees

Actual measurements, derived from reading studies, correspond closely to these estimates. Dodge and Cline[3] cite a mean of 22.9 milliseconds for interfixation movements ranging from two to seven degrees, and Shen[9] reports 10 to 15 milliseconds for a visual angle of slightly more than one degree. Tinker[11] reports 10 milliseconds for an interfixation movement of approximately one degree and 17 milliseconds for movements of approximately two degrees.

TABLE 1.—MEAN EYE-MOVEMENT SPEEDS IN MILLISECONDS FOR VARIOUS ANGLES

	Angles in degrees							
	5	10	15	20	25	30	35	40
Tinker: 7 men	26.6	38.3	48.7	60.6	68.6	78.6	89.2	....
Tinker: 7 women	26.1	36.1	46.6	56.2	63.9	73.7	85.3	....
Tinker: 14 subjects	26.4	37.2	47.7	58.4	66.3	76.2	87.3	....
D and C: 3 subjects	28.8	38.8	48.2	54.8	....	80.4	....	99.9

Examination of Table 1 reveals that with each successive increase of five degrees in visual angle, the time for eye movements was increased by about 10 milliseconds. However, the apparent implication of 10 milliseconds duration for each five degrees of horizontal movement is false. The data show that doubling the visual angle does not double the time taken for eye movements. Thus the time for ten degrees is only forty-one per cent greater than for five degrees; for twenty degrees it is only fifty-seven per cent greater than for ten degrees; and for thirty degrees it is only sixty per cent greater than for fifteen degrees. Dodge[2] suggests that this lack of a one-to-one relationship between amount of movement and time taken for the movement is because the maximum velocity (second phase) of the eye during saccadic movements of large amplitude

is greater than the maximum velocity during movements of small amplitude.

According to the data obtained in various studies, the saccadic reading movements between fixations and to the right along a line of print extend on the average over a visual angle of about one to four degrees and take from about 10 to 23 milliseconds, respectively. In individual cases, the interfixation movements may cover up to seven degrees of visual angle. Return sweeps of the eye to the left from the end of one line to the beginning of the next extend over about twelve to twenty degrees and take about 40 to 54 milliseconds, respectively. The angular extent of saccadic movements to the right along the line of print varies with the comprehension requirements of the reading material. Where the reading is highly analytical, as reading mathematical formulas, the extent of the eye movements is small. True movements, however, seldom cover much less than one degree. For reading familiar prose material, the interfixation moves extend over a larger visual angle. When the extent of the interfixation eye movements is short, there are many more fixational pauses for reading a given amount of material. As noted above, the pauses relative to the moves have long durations. Consequently, the per cent of the reading time taken by the moves will vary with changes in the oculomotor behavior produced by changes in the content or in the comprehension requirements of the material read.

Erdmann and Dodge were the first to point out that a very large portion of the reading time was devoted to pauses. Their data as revised by Tinker[11] showed that on the average 6.4 per cent or one-sixteenth of the reading time was taken by the eye movements. For reading Japanese the eye movement time accounted for four to five and one-half per cent of the reading time. The effect of changes in content of reading material upon time taken for eye movements was first studied by Tinker.[11] Subjects read prose, algebra text, chemistry text and mathematical formulas. The portion of the reading time taken by the eye movements varied as follows: 7.6 per cent for prose; 6.5 per cent for algebra; 6.6 per cent or chemistry; and 2.8 per cent for algebra formulas. The total average showed that 5.9 per cent of the reading time was consumed by the eye movements. Thus there appeared a tendency for a smaller portion

of the reading time to be devoted to movements as the subject was forced to employ a more careful and more analytical type of reading. Examination of the original data reveals that there was a marked increase in both number of fixational pauses and in pause duration for reading the algebra and chemistry in comparison with the prose. This tendency was much more prominent for reading formulas in comparison with the prose. This change in oculomotor behavior as the reading becomes more exacting in terms of attention to facts and relationships appears, therefore, to decrease the relative time devoted to the eye movements, i.e., from 7.6 to 2.8 per cent. There should be more data to check the validity of this trend.

In studying the eye movements of good readers, Walker[15] found that ten per cent of the reading time was devoted to moves and ninety per cent to the pauses.* This value of ten per cent for movement time is the largest reported in the literature. Walker suggests that this lack of correspondence with other published results is accounted for by the high reading rate of his group of subjects. This is possibly true. However, since Walker's readers were highly selected, the result for movement time versus pause time cannot be considered typical of readers in general. Walker further states that the interfixation time cannot be reduced to any great exent as it is chiefly physiological in nature. This, however, would not prevent the ratio between movement time and pause time from changing. As noted above, certain exacting reading situations produce marked increases in number and duration of pauses. It is this kind of change that brings about a variation in the portion of time devoted to the saccadic movements in reading. As a matter of fact, since there are more interfixation movements in analytical reading, the total movement time increases in comparison with the less exacting type of reading. But, since the number and duration of pauses increase even more, the ratio of total movement time to total pause time becomes smaller.

To check previous findings and to obtain a more comprehensive picture of the relation of total eye-movement time to

* Incidently, in quoting Tinker,[11] Walker made an error. He states that Tinker found 10 per cent of reading time devoted to movements and 90 per cent to pauses in reading formulas. Tinker's figures are 2.8 and 97.2 per cent respectively.

total pause-duration time (perception time) a comprehensive investigation was undertaken by Tinker.[12] The purpose was to determine the variation in portion of reading time taken by eye movements with variation in printing arrangement, variation in difficulty of textual material read, and variation in type of response, i.e., oral vs. silent reading.

Eye movements were photographed while ten adult subjects read nine selections of approximately ten lines each. The summary of results is presented in Table 2. The per cent (9.6) of reading time devoted to eye movements for the very easy prose (fourth-grade level) is comparable to the ten per cent cited by Walker[15] for the reading of prose by a group of highly efficient readers. Shifting from silent to oral reading decreases the portion of reading time taken by movements. A similar trend is found in going (a) from medium or long lines to very short lines and (b) from very easy prose to either ordinary narrative prose or hard scientific prose. The per cent (5.3) of reading time taken by eye movements while reading the algebra problems was least of all. For the speed of reading test, the

TABLE 2.—PORTION OF READING TIME TAKEN BY MOVES AND BY PAUSES WHILE READING VARIOUS KINDS OF MATERIAL

Kind of material	Line width: picas	Kind of reading	Move, time in per cent	Pause, time in per cent	Ratio: movement to pause†
Very easy prose.....	25	silent	9.6	90.4	1/9.8
Very easy prose.....	25	oral	6.2	93.8	1/15.7
Easy narrative prose...	9	silent	6.4	93.6	1/15.1
Easy narrative prose...	19	silent	8.1	91.9	1/11.6
Easy narrative prose...	40	silent	8.5	91.5	1/11.1
Hard scientific prose....	25	silent	7.3	92.7	1.13.2
Algebra problems.....	20	silent	5.3	94.7	1/18.8
Speed reading test.....	19	silent*	7.9	92.1	1/11.9
Multiple-c. exam.....	25	silent*	6.2	93.8	1/15.6
Total of all.....	..	.....	7.3	92.7	1/12.7

* Single word oral response given for each item to show comprehension.

† The ratio in the far right column was computed from original data, not from the percentages.

per cent of time devoted to movements was 7.9. This is about the same as for reading easy narrative prose printed in a nineteen-pica line width. In reading the multiple choice examination questions, 6.2 per cent of the reading time was taken by the moves. When all the samples of reading are considered, the average portion of reading time taken by eye movements is 7.3 per cent.

These data reveal certain rather definite trends. In no situation is there more than ten per cent of the reading time devoted to eye movements. In reading situations where comprehension requirements are at a minimum and where perception is smooth and rapid, the portion of reading time devoted to eye movements approaches ten per cent. But in the more complex types of reading, in which the mental processes of apprehension and assimulation become involved, relatively less time is devoted to the movements and more to pauses which are the periods of perception. In oral reading, which for efficient readers is always slower than silent reading, and in reading narrative in very short lines which retards rate of reading, there is also a tendency to devote relatively more time to the pauses (i.e., relatively less time for movements).

There is, therefore, appreciable variation in the per cent of time taken by eye movements in reading in comparison with pause or perception time. This is due to variation in the requirements of the reading situation. In no situation does the movement time exceed ten per cent of the reading time. For most reading situations the portion of time taken by the movements is six to eight per cent of the reading time. This means that the pauses or perception time account for about ninety-four to ninety-two per cent, respectively, of the reading time. The ratio of movement to pause time tends to be, therefore, about one to fifteen and one-half to one to eleven and one-half. In other words, the eyes are motionless a large portion of the time during reading.

SUMMARY AND CONCLUSIONS

1) The purpose of the present report is to coördinate the eye-movement data reported by the author in previous studies in order to bring out the relation of eye-movement time to pause duration in reading.

2) The characteristics of saccadic eye movements are essentially the same whether they are interfixation movements in reading or other excursions from one fixation to another in the visual field.

3) There are significant individual differences in speed of saccadic movements.

4) Although the speed of saccadic eye movements is inversely proportional to the extent of the movement, it is not a one-to-one relationship due to the fact that the maximum velocity of the eye during movements of large amplitude is greater than the maximum velocity during movements of small amplitude.

5) In general, saccadic movements to the right along a line of print extend over a visual angle of from one to four degrees and take from about 10 to 23 milliseconds, respectively. Return sweeps on the average cover about twelve to twenty degrees and take 40 to 54 milliseconds.

6) Fixational pauses, the periods of clear vision during reading, take relatively much more of the reading time than the saccadic movements during which there is no clear vision.

7) The portion of the reading time taken by the eye movements varies with the comprehension demands of the reading situation. In general, the more careful and analytical the reading, the smaller the relative time taken by the eye movements.

8) In no reading situation yet studied is the movement time greater than ten per cent of the reading time. For most reading situations, movements take six to eight per cent of the reading time. Where the processes of apprehension and assimilation become very involved, as in reading mathematical formulas, the relative time taken by the moves may be as low as 2.8 per cent.

9) The changes from about ten per cent of the reading time to only about three per cent for eye movements in going from easy to difficult reading is due to the marked increase in both pause frequency and pause duration required for the more difficult and more complex types of reading.

REFERENCES

1) Benedict, F. G., Miles, W. R., Roth, P., and Smith, H. W. *Human vitality and effciency under prolonged restricted diet.* Washington: Carnegie Institute of Washington, 1919, No. 280.

2) Dodge, R. "Five types of eye movement in the horizontal meridian plane of the field of regard." *Amer. J. Physiol.*, 1903, 8, 307–329.

3) Dodge, R., and Cline, T. S. "The angle velocity of eye movements." *Psychol. Rev.*, 1901, 8, 145–157.

4) Erdmann, B., and Dodge, R. *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage.* Halle, 1898.

5) Huey, E. B. "On the psychology and physiology of reading." *Amer. J. Psychol.*, 1900, 11, 283–302.

6) Lamansky, S. "Über die Winkelgeschwindigkeit der Blickbewegung." *Pflüger's Arch. f. d. ges. Physiol.*, 1869, 2, 418f.

7) Miles, W. R. *Alcohol and human efficiency.* Washington: Carnegie Institute of Washington, 1924, No. 333.

8) Miles, W. R. "Horizontal eye movements at the onset of sleep." *Psychol. Rev.*, 1929, 36, 122–141.

9) Shen, E. "An analysis of eye movements in the reading of Chinese." *J. Exper. Psychol.*, 1927, 10, 158–183.

10) Tinker, M. A. "Apparatus for recording eye movements." *Amer. J. Psychol.*, 1931, 43, 115–118.

11) Tinker, M. A. "Eye movement duration, pause duration, and reading time." *Psychol. Rev.*, 1928, 35, 385–397.

12) Tinker, M. A. "Time taken by eye movements in reading." *J. Genet. Psychol.*, 1936, 48, 468–471.

13) Tinker, M. A. "Individual and sex differences in speed of saccadic eye movements." Chapter xv in *Studies in Personality.* New York: McGraw-Hill Book Co., 1942, 271–280.

14) Travis, R. C. "The latency and velocity of the eye in saccadic movements." *Psychol. Monog.*, 1936, 47, No. 212, 242–249.

15) Walker, R.. Y. "The eye movements of good readers." *Psychol. Monog.* 1933, 4s, 95–1117.

THE RELATIONSHIP BETWEEN SCORES ON THE BELL ADJUSTMENT INVENTORY AND PARTICIPATION IN EXTRACURRICULAR ACTIVITIES

HENRY P. SMITH

School of Education, Syracuse University

INTRODUCTION AND PROBLEM

Presumably the fundamental purposes of extracurricular activity programs in the secondary schools are to promote by means of informal social group participation the development of better social and personal adjustment on the part of students and to stimulate the development of special interests and skills. It would thus seem desirable to collect information concerning the personal adjustment of both those individuals who participate in various types of extracurricular activities and those who participate in none. A previous paper[1] by the writer has already shown that extracurricular activities are highly selective with respect to socio-economic status—in some activities practically all of the participants are drawn from the top half of the school population. Does a similar circumstance obtain with respect to social adjustment? Is it possible that extracurricular activities defeat one of their primary purposes by failing to reach those who most need help?

In the previous paper mentioned above only incidental notice was taken of the relationship found to occur between extracurricular participation and scores on the Bell Adjustment Inventory. Since these latter facts have interesting relevancy to the problem outlined above, they are reported in more detail here. This paper will report data on social, emotional, and home adjustment scores of the Bell Adjustment Inventory from seventeen hundred fifty-one students in the tenth, eleventh, and twelfth grades of a large high school in Iowa. The facts are analyzed to

---

[1] Smith, Henry P., "A Study in the Selective Character of American Secondary Education: Participation in School Activities as Conditioned by Socio-economic Status and Other Factors," *Jour. Edu. Psychol.*, 1945, 36, 229–246.

TABLE I.—MEAN SCORES OF STUDENTS PARTICIPATING IN VARIOUS EXTRA-CURRICULAR ACTIVITIES COMPARED WITH THE SCHOOL MEAN ON THE SOCIAL ADJUSTMENT SECTION OF THE BELL INVENTORY

Activity	N	Sample Mean	Pop. Mean	Diff.	Value of *t*	*t* sig 1% Level
Football	23	10.91	13.17	2.26	1.59	2.82
Track	51	11.63	13.46	1.83	1.83	2.68
Wrestling	27	15.04	13.44	1.60	1.10	2.78
Golf	9	13.22	13.48	.26	.09	3.36
Tennis	7	9.57	13.32	3.75	2.83	3.71
Baseball	22	11.41	13.51	2.10	1.38	2.83
Basketball	19	12.68	13.80	1.12	.59	2.88
Delta S Sor	24	11.42	13.80	2.38	1.97	2.81
Delta U Sor	12	10.42	13.82	3.40	1.84	3.11
Kappa D. Frat	14	8.79	13.20	4.41	2.79	3.01
Delta Th Frat	16	8.31	13.40	5.09	3.01	2.95
Alpha Z Frat	9	7.11	13.09	5.98	2.09	3.36
Spanish Club	100	11.18	13.68	2.50	3.97	2.63
German Club	19	9.11	13.36	4.25	2.59	2.88
French Club	16	13.69	13.63	.06	.03	2.95
Latin Club	91	11.51	13.91	2.40	3.33	2.64
Dramatic Club	22	8.09	13.54	5.45	3.87	2.83
NFL (Debate)	19	6.74	13.08	6.34	6.04	2.88
Art Club	33	12.64	13.85	1.21	.99	2.74
Normal Club	16	10.31	13.72	3.41	2.37	2.95
Jane Addams Cl.	32	14.69	13.74	.95	.83	2.75
Girls Ath Assn	156	12.89	13.86	.95	1.64	2.60
Hi-Y	54	8.87	13.29	4.42	5.70	2.68
Quill Club	36	8.42	13.73	5.31	4.58	2.73
Student Club	167	11.10	13.88	2.78	5.55	2.60
Band-Orchestra	101	12.20	13.39	1.19	1.64	2.63
Voice	184	11.39	13.84	2.45	4.93	2.62
Publications	36	7.86	13.58	5.72	5.00	2.73
Leadership	191	10.59	13.60	3.01	5.95	2.61
ROTC Officers	34	9.41	12.26	2.85	2.24	2.64
Sec'y-Treas	184	11.39	13.84	2.45	4.93	2.62

show relationships to participation in extracurricular activities, to grade level and to sex.

FINDINGS

Table I presents the mean social adjustment scores of students who participated in the thirty-one possible extracurricular activities. It will be noted that in all but three extracurricular activities the members earned a more favorable social adjustment score than did the school population. Twelve of the deviations toward better adjustment are statistically significant at the one-per-cent level, and five more are significant at the

TABLE II.—MEANS AND STANDARD ERRORS OF SCORES ON THE EMOTIONAL, SOCIAL, AND HOME ADJUSTMENT SECTIONS OF THE BELL PERSONALITY INVENTORY BY GRADE AND SEX
(N = 252 for the smallest group)[2]

	Grade	Boys		Girls		Both	
		Mean	σM	Mean	σM	Mean	σM
Emotional	10	7.87	±.34	12.26	±.36	10.32	±.26
	11	7.63	±.36	12.37	±.37	10.22	±.28
	12	7.13	±.31	12.56	±.39	9.85	±.23
	All	7.35	±.19	12.38	±.22	10.14	±.16
Social	10	14.06	±.41	14.06	±.39	14.06	±.29
	11	14.18	±.45	13.39	±.40	13.75	±.30
	12	12.26	±.46	14.03	±.47	13.15	±.33
	All	13.50	±.26	13.83	±.24	13.68	±.18
Home	10	7.07	±.32	8.15	±.34	7.66	±.24
	11	6.99	±.36	7.29	±.34	7.15	±.25
	12	6.32	±.31	7.58	±.35	6.95	±.24
	All	6.80	±.19	7.70	±.20	7.28	±.14

[2] The following differences are significant at the one per cent level: tenth vs twelfth and eleventh vs twelfth grade boys on Social Adjustment; total boys vs total girls on Home Adjustment; twelfth grade boys vs twelfth grade girls on Social Adjustment; and all differences between sexes on Emotional Adjustment.

five-per-cent level. The population means in each case were determined on the basis of the school grade and sex of the persons participating in the activity, with the computed population mean weighted in proportion to the extent to which each activity was composed of members of either sex and drew from the several grades. For example, if ten eleventh-grade and twenty twelfth-grade boys participated in an activity the mean social adjustment score for all eleventh-grade boys was multiplied by ten and the mean score for all twelfth-grade boys was multiplied by twenty and the sum of these products divided by thirty in the process of obtaining an estimate of the population mean. Despite the notable differences apparent in the social adjustment scores, no significant differences were obtained when the mean scores of students participating in the various activities were compared with their population on the basis of emotional and home adjustment.

There was, however, a rather marked sex difference in the emotional adjustment scores and some sex differences in both home and social adjustment scores. In all cases where significant sex differences were found, they were in favor of the boys. These differences are shown in Table II.

No explanation is evident for the sex differences in adjustment scores, but it might be speculated that group mores allow girls to freely admit more fears than do the mores of boys, or that the mores actually promote a more emotionalized reaction pattern. Sex differences in test responses of this type have been noticed elsewhere.

DISCUSSION

Several different although perhaps somewhat related explanations may be advanced for the relationship existing between social adjustment and participation in extracurricular activities:

1) Participation has promoted better social adjustment and/or non-participation has led to maladjustment. If either or both of these possible explanations are accepted, the implications are that the extracurricular activities must be brought more directly into the curriculum in order that the advantages of participation may accrue to all.

Studies of the effect of nurture on test intelligence would lead one to believe that perhaps social adjustment also is determined,

at least in part, by the environment and that more social experience in the home and the school, particularly during the early years, might well lead to better adjustment. Some students in the school studied were prevented from participating because of extraneous factors. Many activities came after the close of the regular school day rather than during a special activity period; others were unable to participate because of expense—certain activities required payment of dues or the purchase of special equipment.

2) Social adjustment is accompanied by a tendency toward participation. Thus, participation in extracurricular activities may be merely another symptom of good social adjustment rather than a causal condition promoting adjustment.

3) The relationship between social adjustment and participation is due to the mutual correlation of these variables with either intelligence or socio-economic status, or both. It is known that some relationship does exist between adjustment and both intelligence and socio-economic status. The author has previously reported a correlation of .37 between Bell Social Adjustment score and socio-economic score as obtained on the Sims Score Card.

Further research is obviously needed to determine which of the foregoing 'explanations' is the major cause of the relationships observed.

It is the writer's guess that the differences between participation and non-participation might be accounted for primarily on the basis of selection (#2 above), either of youngsters of already superior social adjustment or of better economic status, or both. If this guess is correct, it can further be argued that extracurricular activities are not making a great contribution to those who need social development most, and that, while there are real advantages in maintaining the volunteer and informal character of these activities, there is a definite need to capitalize upon their potentialities. Particularly is there need for alert teachers subtly and tactfully to 'ease' a socially-undeveloped child into activities of this type. And there is need for research to determine the effectiveness of such prompted participation on personality development.

Such relationships as are reported here add to our knowledge of the validity of the Bell Inventory. To the extent that a test

is sensitive to differences normally expected between groups it might be considered valid. Thus confidence in the social adjustment section of the Inventory is increased as it is found to differentiate between students who participate in social activities and those who do not. In general, judgements of the validity of a test must be based on the accumulated research involving its use.

A STUDY COMPARING TWO METHODS OF DEVELOPING ART APPRECIATION WITH COLLEGE STUDENTS

EDNA A. BOTTORF

State Teachers College, Lock Haven, Pa.

Educators of today are becoming increasingly aware of a situation which exists in the average community, a situation which deals with the good taste or judgment—or lack of it—of the average individual. It is a situation which is legitimate property of the public schools; if education is to prepare one to live the best life possible for the individual, it must make one acquainted with knowledge of what is the best and how to choose the best from the great array of forms that confront one.

Knowledge of good art forms is one of the most important knowledges needed, since the individual is constantly surrounded by objects and is meeting and selecting new ones. This knowledge is so important, not only because the individual is happier and better with beautiful surroundings, but because he must himself do the choosing of the beautiful and the rejecting of the poorer products. Each individual is a potential consumer of art, but comparatively few become producers of art. Even should it be possible for each person always to have the best supplied to him, he would not enjoy himself and his life so much as if he made his own selections and knew why he did so. Enjoyment is keener for knowing that which is enjoyable.

But can appreciation of art be taught? What seem to be the determining factors in developing it? If it can be taught, the public schools are the logical places for it to be done. They are equipped to supply the necessary knowledge; the methods of disseminating information and developing attitudes are theirs, and the number of people that can be benefitted is larger through the schools than any other channel. It is no wonder that the educators are concerned over the situation.

RELATED STUDIES

Examining the studies carried out by interested individuals, and their conclusions relative to the nature of art appreciation and its development, varying attitudes are found.

Bulley[5] and Drought[12] considered general education to be the largest factor of influence in artistic judgments. In testing the appreciation of their subjects (both ranged from small children to adults) Bulley employed a test comprising objects of household use while Drought employed drawings of varying degrees of quality.

Wall,[39] using a wide range of aesthetic forms including music, literature, etc., in his testing, decided that aesthetic taste was native, intuitive, and universal with development occurring at certain age periods.

In opposition to these, and those others who would claim aesthetic appreciation can be developed through some sort of instruction in art, Bennett[1] says: "Appreciation of art is not the same as a knowledge of art. It is not something absorbed from books, or even from teachers; it is something which must be gained from experience. To know many facts about works of art, to learn the names of artists, to be familiar with events in the lives of painters and sculptors and craftsmen, even to be personally acquainted with artists and with the character of their work, is not to appreciate art. The essence of art appreciation is not primarily a matter of the intellect, but of the emotions."

We do not disagree with Bennett's first statement that appreciation of art is not the same as a knowledge of art. But can the two be kept entirely apart? Would not the one necessarily involve the others? Continuing with Bennett's statement, one might hesitate. His comments do not seem altogether consistent. He claims appreciation must be gained from experience. Would not knowledge about works of art or personal acquaintance with artists and the character of their work be considered 'experience'? His last statement might be meant to distinguish between an intellectual experience—knowledge; and an emotional experience—appreciation. Again, would not this knowledge about works of art and acquaintance with artists and the character of their work involve an 'emotional experience'? Could it be purely an intellectual one?

Jewell, art critic of the *New York Times*, would not agree with Bennett. In a recent article[19] (discussing the critic, Vollard) he says: "We may be sure Vollard knew a great deal about it (painting). And what he knew he had learned in the best possible way, through his own experience; his own observation;

his personal contact with artists and with those who, though not themselves artists, are actuated by a profound love of art."

In contrast to these few cases, where instruction was rejected or neglected as a possible factor in the development of appreciation, there are many more studies where training is wholly, or in part, given the credit.

Christensen and Karwoski,[11] after the testing involved in constructing a test in appreciation (using architecture, design, painting, and industrial art subject-matter), summarized as follows:

"1) The responses of people of various degrees of art training are similar in character and differ principally in the number of art objects appreciated rather than the kind.

"2) There is a wide overlapping between selected groups of people, and especially that the untrained group overlaps into all the other groups.

The first fact implies that art appreciation is a matter of education and varies with the degree of training in art; the second, that original nature determines somewhat one's ability to appreciate art."

They found a fairly low correlation (.299 for girls and .264 for boys) between appreciation scores on their test and the IQ's of those tested. They decided, accordingly, that "art appreciation . . . brings out a special phase of mental activity which is not reached by the intelligence tests."

Knauber[21] says: "Everyone needs art training to help him raise himself from a lower level to a higher one in the appreciation of form and color." This same view is held by Longman.[22]

Pintner,[30] in testing children between the ages of six and fourteen years and university students by using several variations of chosen pictures, concluded: "Aesthetic judgment would seem to grow more uniform as the child grows older. Conformity to conventional standards evidently develops rapidly from the age of nine onwards." Although he does not say whether the growth was due to maturity or training, his phrase, "conformity to conventional standards," would imply training.

Weldele[42] definitely rejected maturity as the cause of growth in appreciation in favor of training. She found that "both elementary- and secondary-school art tended to better the

judgment of pupils in everyday art problems; that maturity without some art training did not better their judgment, . . . and that attending art exhibits improved the art judgment of pupils who attended them."

Pan[28] would emphasize the importance of familiarization in influencing aesthetic judgment, although its effect was found to be lost in the course of time. This loss may have been due to the very short period in which familiarization was allowed to occur in this study.

Studies in which the effect of instruction upon appreciation was measured have been conducted by Waymack,[41] White and Johnson[44] and Carroll.[9] Larger scores were found after instruction by both Waymack and by White and Johnson in picture appreciation by grade-school children. Carroll found those having more art instruction had higher scores, although among his college students he did not consider them as significantly higher. The scores were secured from results on the Meier-Seashore (using pictures) and the McAdory (using pictures, industrial art, etc.) tests.

Norton[27] also used the Meier-Seashore test to check the effect of one year's customary classroom art work on the aesthetic judgment of seventh-grade pupils and found a positive improvement taking place. She also found that intelligence and mental age have a greater influence in promoting this growth than chronological age. This would strengthen the inference that Pintner's study referred to training rather than maturity, but it would be at variance with Christensen's and Karwoski's conclusions that intelligence had little effect on developing appreciation. It might be noted, despite Christensen's statement to this effect, that the r's he found were positive (.264 and .299).

Voss[38] came to the same conclusions as Norton where the principles of art were studied with practice by elementary-school children. She found a significant increase in aesthetic judgment and analysis abilities without necessarily an increase in production ability.

Sutherland,[34] using the Meier-Seashore test for checking, concluded that significant differences in the amount of art appreciation were found only when the group was divided on the basis of the nature of the art work taken. A significant

advantage was found for design principles as contrasted with all other types of work. These findings might be considered to give a clue to the best way to develop aesthetic appreciation.

Boswell[2] and Longman[22] would include both participation in technical processes and formal study as necessary for this development. Boswell claims that art appreciation is a personal, emotional, and intellectual reaction. Several others also stress the emotional part of the aesthetic experience: Bennett and Mursell[26] would give it particular importance. This does not necessarily imply an emotional instability. Carroll[8] and Druley[13] found their subjects who rated highest in art also exhibited emotional stability as measured by such tests as the Bernreuter Personality Inventory.

The different types of instruction mentioned have been in general of two kinds; one in which technical skills (Christensen and Karwoski, Weldele, Carroll, Norton, Voss, Sutherland) were stressed, and the other in which some phase of the fine arts was given formal study (Pan, Waymack, White and Johnson). Boswell and Longman suggested both.

Terry[36] presented an historical treatment of the comparative merits of the two attitudes, based largely on library research. Scott[31] reports Speer[32] and others as opposed to execution as a method of developing appreciation: "Teaching through execution, even though the aim be 'appreciation' does not result in such growth in appreciation of art as gained when appreciation is taught 'through class discussions and the making of discriminating judgments' without resort to production or creation. The average class level in art appreciation can be raised higher through methods which aim solely at appreciation. . . . Other functions than intelligence enter into the determination of status in appreciation; there is no relationship existing between intelligence and art judgment and aesthetic sensitivity to things artistic; art appreciation is determined by a combination of native capacity and training."

Scott herself compared two methods of developing appreciation; in one, using the conventional historial approach, in the other emphasizing the cultivation of a perceptive power. The latter method showed a slightly higher gain in appreciation in the small class of college women with whom she experimented.

In the belief that aesthetic judgment can be increased through

a formal study of the fine arts, many colleges have been increasing the number of their courses dealing with a study of historic art. The great art periods and their products are discussed in various ways. The artists themselves—their lives as well as their works—are studied. The historical is the usual approach. It is thought that contact with and knowledge of these beautiful forms will produce a critical appreciation of other and newer art forms. This belief is strengthened by the findings in the studies of Boswell, Scott, Speer, Jewell, Weldele, Pan, Waymack, and White and Johnson, but it is opposed by Bennett.

On the other hand, the public school art program throughout the grades is concerned with creation, production and execution—not to develop these abilities alone, but to develop a critical judgment on the part of the pupils. If execution does not develop appreciation, the work of the schools would appear partly wasted. To substantiate their claims we have the testimony of Boswell, Christensen and Karwoski, Weldele, Carroll, Norton, Voss and Sutherland. Directly opposed to this belief are Speer, Voss, and Meier and Bulley.[4] Supporting the belief are those who use the Meier-Seashore and McAdory Art Tests not only as measures of art ability but also of art judgment. These tests are used indiscriminately for both. Although Meier claims his test is a measure of judgment but not of training, he offers as validation for the test that those with more art training make higher scores.* Carroll[4] reported negligible correlation between scores on these tests and teacher grades, yet he and Eurich[15] found college art majors excelling the general students in their scores. Wallis[40] found r's of .53 and .58, respectively, between scores on the Lewerenz Art Test (designed to measure art ability) and the Meier-Seashore and the McAdory Art tests. Between the scores secured on the last two tests, the r was only .37 in Wallis' study. Meier,[24] author of the Meier-Seashore test, based his test on the belief that art judgment is a reliable measure of art talent.†

Because of the conflicting beliefs about the value of formal

* Swartz[35] believes the Meier-Seashore test measures the aesthetic sense, but that it is not wholly reliable in prognosticating success in art.

† Knauber[21] on the other hand thought they were distinct abilities and planned separate tests for each.

study of the fine arts and of execution in developing aesthetic appreciation, this study was planned to compare these two methods in this respect and to evaluate or examine any ramifications growing out of it.

Accordingly, the first purpose of this study was to discover whether appreciation of art products is acquired better through a study of the history and products of the great art periods or through an individual's practical experience with materials in creative problems. Other purposes would be to discover the relation of intelligence or other factors in developing appreciation with either method.

THE METHODS OF INSTRUCTION

The method of studying Art History was used with a college group of one hundred fifteen. The course extended through one semester and consisted of approximately fifty hours. The first few hours were spent in explaining the principles* underlying all works of art (whether among the so-called fine arts or minor arts and crafts) and in discussing how these principles were applied in actual objects such as, the school building, the classroom, small sculptures, pottery, candleholders, hardware, science tools, etc. The remainder of the time was spent in lectures and discussions, supplemented with illustrative material on the great art periods and their products and the outstanding painters and their pictures.

These principles, which were also taught to the second group as basic to their work, were classified under three headings. The first, called the principles of design, included balance, repetition, variety, opposition, transition, proportion, dominance, subordination, and unity, and were studied as they related both to structural and surface design and their relation to each other. The second classification dealt with the material of which an object is made, its limitations and possibilities. The third was concerned with the purpose or use to which the object was to be put and the needs of such a purpose. The close

---

* It will be recalled that Sutherland claimed a study of design principles held a significant advantage over all other methods in increasing appreciation. Accordingly these principles were stressed with both groups used in this study.

relationship existing between the three—their bearing on each other—was strongly stressed in both groups.

The second group, consisting of twenty pupils, was given instruction for one semester (or approximately seventy hours) in actual designing, construction and decoration of various objects such as imaginative animals or toys, table hot-pads, letter and art folios, cardboard and wooden boxes, block printed greeting cards, pottery, etc. As large a variety of materials as possible was used including paper, cardboard, wood, composition board, dyes, water color, tempera paint, oil paint, printer's ink, gesso, varnish, shellac, soap, clay, and linoleum. Different techniques used included pasting, painting, modeling, carving, and dyeing.

In order to measure the effect of each method in developing appreciation it was decided to administer both before and after the semester's instruction a test involving judgments in appreciation of many types of problems and use the differences found in the two sets of scores as indicative of the influence of the method employed.

CONSTRUCTION OF THE TEST

A search of studies to find a suitable test disclosed a variety of testing materials. Some investigators used pictures selected by themselves or art teachers; some used colored triangles or other art forms. Gordon[18] used pictures of oriental rugs. Although she called them 'fine rugs,' she also claimed they "appeared to have a fairly wide range of aesthetic value," Cattell, Gloscock, and Washburn[10] used thirty-six cosmos prints in black and white of wide range "from the sentimentally popular to the technically great." The best known and most widely used test employing pictures is the Meier-Seashore Art Judgment Test. But with emphasis in the present study upon execution and creation with one group and the investigator's belief that appreciation applies—or should—to the minor arts as well as the fine arts, none of these tests seemed pertinent.

A few tests met some of the requirements. The Appreciation Test by William G. Whitford,*[45] one of the first attempts to use scientific measurement in connection with art work, was made up of a series of problems in design and color. The Test

* Test not available.

of Appreciation of Art by E. O. Christensen*[11] consists of a series of four black and white variations of selected examples of furniture, automobiles, costumes, landscapes, silverware, posters, monograms, design, etc. The subject taking the test is to choose the pleasing example in each case. By pleasing is meant "the artistic quality of the designs which are to be judged regardless of any practical use or application." This limitation was in direct opposition to one of the fundamental attitudes in this particular study—that the artistic worth of an object is in part determined by fitness to use and to material. The Art Test prepared by Margaret McAdory†[23] consisted of a series of four variations of the chosen examples, involving differences in shape and line arrangement, in value of dark and light, and in color schemes. The examples used were furniture, utensils, textiles, clothing, architecture, painting, and objects of other plastic and graphic arts. The limitations of reasons for choices (differences in shape and line arrangement, in value of dark and light, and in color schemes) and cost of the test prevented its use.

Since no suitable test was available it was decided to construct a test for the purpose. Wentworth[43] evaluated the Meier-Seashore, McAdory art test, and a test constructed on an actual room, by giving them to three hundred seventy-two college students ranging from those with no art training to those who had had four courses in related art. She found that it is possible to construct a reliable and valid objective test of art appreciation. Accordingly in this problem it was felt defensible to construct a test to meet the situation.

For the test to be used in this study, clippings of photographs and illustrations were chosen from current magazines, newspapers, and books of art on as wide a range of subjects as was possible, (taking into consideration the kind of forms the average individual would meet with most constantly, the kind in which he should show some interest, and the amount of time that would be required to take such a test).

---

* Test not on the market.

† Miss McAdory's test is criticised by Tyler[37] as a test of the recognition of objects but not of enjoyment, although farther on in the article he seems to contradict himself when he says the test will be useful in providing a measure for the development of pupil's judgment of artistic merit.

Fifty-seven forms were finally selected dealing with newspapers, books, costume, objects in the home, architecture, painting, sculpture, and many other forms.

Both good and poor examples were gathered. The final clippings included: (1) forms which were considered fine examples of their type; (2) examples which had very good qualities but also exhibited some poor qualities; and (3) examples which violated the principles of art to a great extent or were on the whole mediocre. From these clippings were chosen six examples for each form to be tested: two of the best, two of the intermediate, and two of the poor types.

The ratings were determined by the investigator and the opinion of authorities as found in their writings. Typical of these ratings were the following examples: (a) in room arrangement, two very simple, pleasing arrangements conforming to the structure of the room—A value; two very crowded and intricate arrangements—C value; and two arrangements partly pleasing and partly confusing—B value; (b) in lettering types, examples of fine Roman and Old English lettering—A value; modern lettering bordering on the 'tricky'—B value; and lettering plainly 'tricky' and hard to read—C value; (c) in architecture, a photograph of Notre Dame, Paris, and the Parthenon—A value; a New England farmhouse and colonial home—B value; and one ornate, pretentious house of the late Nineteenth Century and one plain box-like structure—C value.

RELIABILITY AND VALIDITY

The reliability and validity of this test were then found by giving it to several hundred beginning college students.

Reliability refers to the consistency with which a test performs its intended function. Several methods of finding reliability are used; for instance, repetition of a test at a proper interval, comparison of two comparable forms of a test, or the split-half technique (comparison of the odd and even answers in a test). The latter form was used in this instance. The Pearson Product r between the two halves is .59, the r of the whole test (using the Spearman-Brown formula) is .74, and the index of reliability is .86. These findings compare favorably with similar information from the standardized tests previously

discussed. The reliability coefficient of this particular test is high enough to assure a fair measure of reliability.

The validity of a test refers to its ability to measure the particular thing which it is supposed to measure. The validity of this test was determined by giving the test to a total of two hundred fifty-six college students, one hundred ninety-five of whom had had no college art, and to two adults—one a college art teacher, the other a salesman for an art supply company who had previously taught art in the public schools. The means and median scores for each group, subdivided into sections according to the amount of art training reported on a Personal Inventory Sheets, are found in Table I.

TABLE I.—AVERAGE SCORES MADE BY EACH GROUP (NO COLLEGE ART) ACCORDING TO THE AMOUNT OF ART TRAINING REPORTED

Group	Number in Group	Median Score	Mean	SD
No Art in Grades.....	35	127.8	128.15	12.10
Art in First Six Grades.....	20	132.0	133.25	16.05
Art in First Nine Grades.....	120	135.75	137.33	17.20
Art in Twelve Grades.....	18	137.0	141.45	26.90
Art in Twelve Grades and College.....	4	197.0	198.25	49.04

Table I shows a progressive increase in scores according to the previous art instruction reported. The intervals, it will be noted, are not of equal length but follow the usual school divisions of elementary, junior high, senior high, and college years. The omission of the 'Art in First Nine Grades' would remedy this situation and show even more definite progression from the 'Art in First Six Grades' to 'Art in Twelve Grades.' The presence of this score does not lessen the implications, so it was not omitted.

In Table II, the average scores secured by the sixty-one individuals who had had at least two (but not more than four) semesters of college art are arranged in a similar manner according to amount of public school art reported.

Again a progressive improvement is shown with the differences between groups larger than in Table I. It will be noted that

the averages for the group with no grade art training in this group is higher than the averages for the group having the highest (twelve) grades training in art but no college art instruction. Likewise, the averages for the total group (excluding

TABLE II.—AVERAGE SCORES MADE BY EACH GROUP (HAVING AT LEAST TWO SEMESTERS OF COLLEGE ART) ACCORDING TO REPORTED AMOUNT OF ART TRAINING IN THE GRADE SCHOOLS

Group	Number in Group	Median Score	Mean	SD
No Art in Grades	9	147.0	146.22	8.60
Art in First Six Grades	6	159.5	159.50	19.86
Art in First Nine Grades	42	162.0	165.57	19.50
Art in Twelve Grades	4	197.0	198.25	49.04
Total	61	159.8	169.4	24.60
Adult (Art Teachers)	2		244.0	

the instructors) with college art training is higher than those of any group not having college instruction. The two adults (instructors) have a much higher average score than any group.

The data were arranged graphically in six curves which, for want of space, are not reproduced here. The scores were arranged vertically and the frequencies horizontally. Figure 1 covered the data for the group having no college art arranged according to the reported art training in the grades; Figure 2, the data for the groups—arranged according to the art training reported in the grades—which have had two to four semesters of college art. The data for those having art in the twelve grades and college, too, were thus recorded twice. Figure 1 and Figure 2 were arranged beside each other in order to show the strong progression of increase in scores, and it was noted that the profiles of the groups having no grade training in art paralleled quite closely the profiles of the groups reporting art training in the first six grades, both for those having no college art training and for those who did. The same situation is true for those reporting nine grades and those reporting twelve grades of art training—particularly for those not having college art. This situation, when explained, increases the validity

of the test rather than decreases it. Many of the pupils reporting art training in the first six grades were not sure how many years of training they had and actually had no real art instruction at all. Tracing and coloring of patterns and forms of 'busy work' were called art. This has been found to be true through questioning the pupils as to the nature of their art work. Some of those reporting training in the high school were later found to have reported undirected copying of pictures etc. as training. Rather than decreasing the validity of the test, this information increases it. If such poor, undirected 'busy work'—art by courtesy—increases the scores, real art training would be expected to do even better. Those reporting training in the twelve grades were all (with the exception of the lowest score) good art students in college. All were very much interested in the subject, as evidenced by their taking it their three years in high school where it was an elective subject. The increase in their mean score over the similar group without college art shows to a certain extent how much better the scores might have been with a more accurate and reliable criterion for judging.

Several interesting facts stand out. The mean score for all those having college art is found to be exceeded by only six subjects in the entire group having had no college art. This would definitely show that training does bring about an increase in the scores. It was interesting to note that the lowest scores for each group having no college art is the same, with one exception, 100–104. The lowest scores for each group having college art is also the same, 130–134, but higher than for the 'no college art' group. On the other hand, each subgroup shows a progressively larger 'highest' score according to the amount of art training. In comparing the scores of those having no college art with the mean scores of the groups having college art it was found that a small percentage of those without college art exceed those mean scores. This again would tend to strengthen the validity rating of the test. The mean of the two adults is exceeded by only one college pupil having college art and by no one in the group having had no college art. The one adult score, 176, was made by the individual who had not been teaching art for some years. It is exceeded by only two individuals who had no college art and by a slightly larger number who had college art.

Meier, in part, validates his test by showing the low correlations that were found between IQ scores and scores secured on his test. He did this since so many writers and investigators have claimed that art appreciation does not seem to be a factor needed on the usual intelligence test. The two criteria have been found to correlate around zero by many investigators. This situation was found by Meier,[24] Christensen,[11] Stinson,[33] Carroll,[7,14] Speer,[32] and the present investigator in a previous study.* On the other hand, Garrells[16] and Norton[27] claimed a positive correlation between appreciation and intelligence. The data used in this study showed a correlation of −.001 between IQ's and the initial scores on the test and an r of .08 between IQ's and the increase between the two applications of the test with the control group.

The correlation found between scores and art training received in the grades was .32 for those having no college art and .41 for those having college art. The upper limit of the test's validity from this information would be .43 for those not having college art and .56 for those having college art.

Considering the validity of the test from these various angles, and its reliability as previously shown, the scores secured on this test seemed a legitimate evaluation of ability to judge art objects and a practical test to use in this study.

SECURING THE DATA

As has been noted in previous pages, different investigators claimed factors other than training were likely to have a part in determining the progress in art appreciation of the individual. Some of these factors can be measured, while for others there are as yet no valid measuring instruments.

No one would doubt that appreciation of art is accompanied in part by an emotional reaction to the object. This can not be measured by the emotional stability of the individual, as was found by Carroll and Druley, who subjected their pupils to the Bernreuter Personality Inventory and other tests to measure emotional stability. Those doing well in art were found to be emotionally stable.

---

* *A Study Comparing Art Abilities and General Intelligence of College Students.*

General education as claimed by Bulley and Drought was not a factor in the present study, since all pupils used were college freshmen. Neither was chronological age or maturity, stressed by Wall and Pintner, but rejected by Weldele, Miller, and Norton, a factor to be considered since the age of the pupils used varied little.

Jewell stressed contact with art; Pan, familiarization; and Weldele, museum visits as influencing appreciation. Although visits to museums were to be checked by the pupils on their Personal Inventory sheets, the data could not be used since some of the so-called museums visited were local historical ones and pupils expressed doubt as to the number of visits. Travel—through which contact with art might have been increased—might also be a factor in developing appreciation. No one appears to have studied this point. To measure the effect of travel would be difficult, since number of miles, number and nature of places visited, etc., would not be reliable measuring devices in themselves.

The effect of intelligence as measured by intelligence tests is disputed by investigators, yet, on the whole, even those who deny its effect usually found positive, if very low, correlations between intelligence and art appreciation. According the IQ's of the pupils used in this study were secured from high-school records and from the Otis Self-Administering test and were used as a factor in comparing groups and individuals.

TABLE III.—CORRELATIONS BETWEEN YEARS OF ART TRAINING AND OTHER FACTORS USED IN EQUATING GROUPS

Factors Used in Correlations	*r*
Intelligence Quotients	.09
Art Grades	.10
Initial Test Scores	.12
Gains of End over Initial Test Scores	−.008

$N = 116$.

Positive, if fairly low, correlations have also been found between art grades and scores on art appreciation tests. Carroll[14,9] and Eurich[14] found an r of .14 ± .07 between art grades of college students and their scores on the McAdory Art Test and an r of .40 ± .06 between the art grades of this same group and their scores on the Meier-Seashore Art Judgment Test.

If instruction is to influence the development of appreciation (an assumption basic to this study), the art grades of the pupils should show to some extent the progress of the individual or, at least, be a measure of the effort he expended in the course, and thus be a valid criterion for equating groups.

Years of previous art training was rejected as a factor in influencing appreciation since the data collected appeared too unreliable. Correlations computed between the years of art training and other factors selected for this study were so low as to warrant their rejection. (See Table III.)

The low correlation between years of art training and initial test scores (.12) is lower with this group than that secured with the same criteria when proving the validation of the test. This drop may have several explanations. The most important—besides the weaknesses inherent in the 'years of art' criterion already explained—is the fact that in this case the individual number of years reported were added together, no matter when taken, to secure the data. In the previous case only those reporting art training in consecutive years and for definite periods throughout the elementary grades; throughout the elementary grades and junior high school; throughout the elementary grades and junior and senior high school; and throughout the grades I to XII and college were counted. The unreliability of the first method here to the second is apparent. For instance, a pupil having art training in grades I, II, and III would be listed with one having art instruction in grades X, XI, and XII. The difference in instruction, the recency of the time when taken, and other factors would affect the scores of these individuals to a great extent, yet such effects would be discounted and thus distort results. The lowness of the r's (Table III) themselves would have little effect on the results as worked out, so that the rejection of this criterion as an equating factor is of no importance to the problem.

Accordingly, the criteria used in the experiment were: (1) IQ's representative of the equipment the pupil could bring to the situation; (2) art grades, representative of the effort actually expended on the situation; (3) initial art judgment test scores, representative of the amount of appreciation with which the pupil began the experiment; and (4) the gain—difference between initial and end scores on the art judgment test—representa-

tive of the improvement brought about during the period of experimentation.

TREATMENT OF THE DATA

There were one hundred fifteen individuals in the group who studied the principles of art through the history of art and only twenty who studied and applied the principles of art to actual objects or situations. It was almost impossible to match individuals according to the factors decided upon: IQ's, art grades, and initial art scores. Accordingly, it was decided to use the regression technique suggested by Peters[29] for the matching of groups, which disposes of the necessity for having precisely matched pairs.

According to this method, scores are predicted for hypothetical pupils having the same equating factors as the members of the experimental group but supposedly reacting as the pupils of the control group reacted. The amount of influence of each equating factor in determining the criterion—that factor which showed the effect of the method employed (in this case gains in the appreciation test)—is found numerically and applied to the net set of equating factors to predict a score such a set would have made under similar conditions. In effect this method provides results for two groups identical in equating factors but each working under a different method—the one result secured from an actual situation, the other secured from a hypothetical situation. Since, in this technique, beta weightings for each criterion are secured from correlations between equating factors and correlations secured from larger populations are more reliable than those secured from smaller populations, it was decided to use the larger group—the 'history of art' group—as the control group.

The following zero-order correlations were obtained among the three equating factors (IQ represented by subscript 1, art grades by subscript 2, and initial art test scores by subscript 3) and the criterion (gains in appreciation) represented by subscript 0:

$r_{01} = .08 \pm .06$ $\quad r_{12} = .36 \pm .05$ $\quad r_{23} = .15 \pm .06$

$r_{02} = .32 \pm .06$ $\quad r_{13} = -.001 \pm .06$

$r_{03} = -.43 \pm .05$

Using these r's, the following beta weightings for these factors were secured:*

$$\text{Beta}_1 = -.0716$$
$$\text{Beta}_2 = .4197$$
$$\text{Beta}_3 = -.493$$

Since the regression technique required the means and standard deviations of the factors used, they were computed as follows:†

	Means	SD
1. Intelligence Quotients	108.81	10.09
2. Art Grades	5.67	1.67
3. Initial Art Scores	130.43	14.57
4. Gains of end test scores over initial test scores	10.01	14.07

Using these values computed from the data of the control group, gains were predicted for the experimental group using the formula from Peters and Van Voorhis.‡

Table IV records these predicted scores.

In Table IV the scores made by the experimental group in the equating factors are followed by the gains actually made by these pupils, the gains predicted for them using the regression technique, and the differences found between actual gains and the predicted gains. The differences show a mean gain of 1.78 for the experimental group over the control group. In order to find out whether this gain was statistically significant-whether the experimental method would always prove better in developing appreciation, as measured by this test, than the control method—the standard error of this score was determined from formulas worked out by Peters.

---

* The betas were secured by use of equations taken from Peters and Van Voorhis ([46], p. 210).

† Art grades were averaged by assigning a numerical value to each grade: A = 11, A− = 10, B+ = 9, B = 8, B− = 7, C+ = 6, C = 5, C− = 4, D+ = 3, D = 2, D− = 1.

‡ Peters and Van Voorhis, p. 199:

$$X_0 = \sigma_0\left(B_{01.23}\frac{X_1}{\sigma_1} + B_{02.13}\frac{X_2}{\sigma_2} + B_{03.12}\frac{X_3}{\sigma_3}\right) - \sigma_0\left(B_{01.23}\frac{M_1}{\sigma_1} + B_{02.13}\frac{M_2}{\sigma_2} + B_{03.12}\frac{M_3}{\sigma_3}\right) + M_0$$

From these formulae the standard error of the differences was found to be 4.32. The ratio of this SE to the difference (1.78) is only .41, which indicates about a two to one chance

TABLE IV.—PREDICTED SCORES IN COMPARISON WITH ATTAINED SCORES FOR THE EXPERIMENTAL GROUP

Pupil	IQ	Art Grade	Initial Test Score	Gains	Predicted Gains	Difference
1	117	B−	128	− 3	+ 15.05	− 18.05
2	106	C+	133	19	+ 10.24	+ 8.76
3	112	B−	140	− 9	+ 9.84	− 18.84
4	113	C	176	−10	− 14.39	+ 4.39
5	105	D	148	10	− 11.01	+ 21.01
6	109	B	175	9	− 2.99	+ 11.99
7	113	D+	153	− 6	− 10.59	+ 4.59
8	125	B−	137	30	+ 9.95	+ 20.05
9	105	B+	154	11	+ 10.94	+ .06
10	109	B−	153	8	+ 3.96	+ 4.04
11	107	B−	128	24	+ 16.00	+ 7.99
12	113	C+	178	− 3	− 11.85	+ 8.85
13	99	C−	132	− 8	+ 4.53	− 12.53
14	110	B−	148	7	+ 6.30	+ .70
15	85	C	137	− 5	+ 6.86	− 11.86
16	90	C	149	−24	+ .67	− 24.67
17	102	C	102	26	+ 21.77	+ 4.23
18	104	B−	147	7	+ 7.28	− .28
19	121	B−	165	17	− 2.99	+ 19.99
20	110	C+	133	15	+ 9.81	+ 5.19
				183	+133.20	+121.84
Total				−68	− 53.80	− 86.23
				115	+ 79.39	+ 35.61
Mean	107.75	6.0	145.8	5.75	+ 3.97	+ 1.78
SD	9.07	1.61	18.14	13.65		12.29
SE of Difference =						4.32
Ratio of Difference to SE of Difference =						.41

Odds against reversal are 65.91 to 34.09 or 1.9 to 1.

against reversal. Statistically this is not enough to warrant any advantages being claimed for the experimental method. From a study of the differences it was noted that fourteen of the

twenty members—or seventy per cent—of the experimental group exceeded their hypothetical mates (or the predicted scores). This large percentage should have some weight in comparing the value of this method with the other method. Should both be about equal in developing appreciation as this data appears to show statistically, other values of each method might be measured and compared, for instance, the value of the skills developed in the 'execution' method as compared with the knowledges acquired through the 'history of art' method. This could not be done in the present study.

The results as found here do not agree with those of Speer and others that appreciation is developed better through a study of the arts than through execution. Their conclusions may have been based on inadequate data. A glance at the mean scores for each group—the mean gains 10.01 for the control group and only 5.75 for the experimental group—would mislead one into forming conclusions similar to those found by these investigators. Groups must be perfectly equated and on more than one criterion in order to form just conclusions. Slight differences in equating factors could produce greater discrepancies in results than might be expected.

SEX DIFFERENCES EVIDENT IN THE STUDY

Since the data secured from the two groups included both girls and boys, it was thought of interest to find out if any sex differences were apparent. Only the data from the control group was used here, since the experimental group was largely girls. The control group was almost equally divided—there were sixty boys and fifty-five girls.

Differences in equating factors and criterion are presented first.

Table V presents the data relative to gains. It is seen that, although the boys showed more variability in their scores, with both advancing and receding gains larger than the girls, the mean gains for the two groups are very close—only .04 of a point different and favoring the boys. The cause for the greater range of scores for the boys is not apparent here.

Table VI shows that the mean IQ of the girls (111.00) was somewhat higher than that of the boys (107.75). Since the beta weighting of IQ's in connection with gains was low, this

TABLE V.—COMPARISON OF THE GAINS OF THE BOYS AND GIRLS IN THE CONTROL GROUP

Gains	Boys		Girls	
	Number of	Per cent of	Number of	Per cent of
45–54	2	3.3		
35–44	2	3.3	1	1.8
25–34	9	15.0	4	7.2
15–24	9	15.0	11	20.0
5–14	20	33.3	15	27.2
– 5–4	12	20.0	15	27.2
–15––6	2	3.3	9	16.3
–25––16	3	5.0		
–35––26	1	1.6		
Mean Gain		9.7		9.3

TABLE VI.—COMPARISON OF THE IQ'S OF THE BOYS AND GIRLS IN THE CONTROL GROUP

IQ's	Boys		Girls	
	Number of	Per cent of	Number of	Per cent of
125–129	3	5.0	4	7.2
120–124	6	10.0	4	7.2
115–119	8	13.3	10	18.2
110–114	7	11.7	11	20.0
105–109	11	18.3	12	21.8
100–104	10	16.6	5	9.1
95– 99	11	18.3	5	9.1
90– 94	4	6.6	1	1.8
85– 89	0	0.	3	5.4
Mean IQ		107.75		111.00

could not have caused the higher gains—nor lower—among the boys.

Table VII presents the data relative to the art grades of the boys and girls.

TABLE VII.—COMPARISON OF THE ART GRADES OF THE BOYS AND GIRLS IN THE CONTROL GROUP

Art Grades	Boys		Girls	
	Number of	Per cent of	Number of	Per cent of
A−	1	1.6		
B+	3	5.0	3	5.4
B	5	8.3	6	10.9
B−	7	11.7	11	20.0
C+	12	20.0	10	18.2
C	18	25.0	9	16.3
C−	7	11.7	11	20.0
D+	6	10.0	5	9.1
D	1	1.6	0	0.
Mean Art Grades	C+ minus .4		C+ minus .25	

A situation similar to that of gains is found here, the boys show a wider range in scores, receiving both higher and lower grades, yet the mean grades for both groups are very close, only .15 of a point, in favor of the girls. Since the beta weighting of art grades in affecting scores was fairly high (.4197), this may have been the factor in causing the wider range in gains among the boys.

In initial test scores, the girls definitely exceeded the boys, the range of their scores starting in the interval above the lowest for the boys and extending to one interval beyond the boys; the mean score for the girls is an interval (almost ten points) higher than the boys. The suggestion of some investigators that girls show a higher degree of aesthetic sensitivity is borne out by this data.* Since the beta weighting of initial test

* See Eurich and Carroll.[14]

score was fairly high and negative—the higher the initial test score, the smaller the amount of gain and the lower the test score the greater the amount of gain*—this would help to explain the slightly higher gains of the boys.

The correlations of these factors for boys and girls revealed interesting information. The boys showed a slightly higher correlation between IQ's and gains (.05), but the standard error of this difference was only .18, giving a ratio of .28 to the difference or 1.6 chances to 1 against reversal. The fact noticeable of very low correlations between IQ's and gains for both boys and girls, strengthens the supposition that this low correlation was somehow connected with the method of study. The r in the same situation—IQ's and gains—in the 'execution' method was definitely greater (.39), although the group was composed mostly of girls. The higher correlation was therefore not due to sex differences but rather to the method of study (or some other factor not brought out in the data).

The correlation between art grades and gains was much higher (.38) for the boys than for the girls. This difference is quite reliable and statistically significant since it has a standard error of .16 which gives a ratio of 2.38 to the difference, representing odds of about 114 to 1 against reversal. This result and the fact that the correlation coefficient was itself fairly large (.51) between the art grades and gains for the boys strengthens the inference that the wider range of art grades among the boys was responsible for a corresponding wider range in their gains. The experimental group (almost all girls), in this instance, agreed more closely with the control group as a whole (r = .35 for the experimental group, r = .32 for the control group) than with either the boys or the girls. These data needs further study and additional support. The experimental group is too small in numbers for complete assurance of any statement. The r in this instance, also, is not quite three times its PE, making it statistically unreliable. The same is true of the r for the girls in the control group (.13 ± .09), while the r's for the whole control group (.32 ± .06) and for the boys of that group (.51 ± .06) are both statistically reliable.

* A situation to be expected. The higher score an individual secured the less chance there would be for improvement since he would constantly be approaching the limit of improvement.

The initial test scores and gains correlated almost the same for both boys and girls with only a difference of .08. This difference has a standard error of .15, which gives a ratio of .53 to the difference or chances of only 2.4 to 1 that this situation will not be reversed. The r's for all groups—including the experimental group—in this situation (initial test scores and gains) are all very close and therefore indicate a consistent tendency for initial test scores and gains to correlate fairly closely (around −.4), and negatively (since all are negative correlations) under different conditions and with both sexes.

Accordingly, it is seen that, with the exception of art grades and gains, the findings in regard to the sexes substantiates that found for the group as a whole and therefore strengthens any conclusions that may be drawn from them.

CONCLUSIONS

The specific conclusions to be drawn from this study are quite clear:

1) Although the 'execution' method was not found to be statistically more productive of aesthetic appreciation than the 'history of art' method, the tendency in that direction appeared so persistent as to strengthen any evidence toward such a conclusion.

2) Intelligence was an important factor in developing appreciation with the 'execution' method but of practically no influence with the 'history of art' method. The 'execution' method appears to be a more inclusive method.

3) Improvement in aesthetic appreciation correlated almost equally well under both methods with the effort put forth by the pupils as measured by their art grades.

4) The initial test scores correlated negatively (around −.4) and practically the same under both methods with gains in appreciation and indicated that the higher the initial score, the smaller the improvement in aesthetic appreciation; and the lower the initial score, the greater the improvement, no matter which method is used.

5) The girls showed a higher degree of aesthetic appreciation than did the boys.

6) The boys appeared more variable than the girls in their art

study and in their development in aesthetic appreciation; they made both higher and lower scores in gains than the girls, corresponding to a similar situation in art grades although the mean scores in both cases were almost identical for both sexes.

7) Under the 'history of art' method, the boys showed a greater tendency to improve in appreciation according to their efforts in their art study as measured by their grades than did the girls.

Stated briefly, we may conclude:

The degree of development of aesthetic appreciation is dependent upon different factors and favors slightly the 'execution' method. Although intelligence does not appear to be one of these factors, its influence depends upon the method of study, being quite strong when combined with the 'execution' method. Effort expended in study (art grades) and the original capacity to appreciate (initial test scores) also exert much influence under both methods, the first in a positive manner, the second negatively. Sex does not affect these factors except in the case of art grades where it increased the effect of that factor in the case of the boys and reduced it for the girls. The girls show a higher degree of aesthetic sensitivity while the boys show more variability in their art ability or effort and in appreciation of the artistic worth of everyday objects.

BIBLIOGRAPHY

1) Bennett, C. A., *Art Training for Life and Industry*, Peoria, Ill.: Manual Arts Press, 1923, p. 61.

2) Boswell, Mary Kathryn, *A Compilation of Theories Relative to Engendering Art Appreciation*, Master's Thesis, George Peabody College for Teachers, Nashville, Tenn., 1930.

3) Bouffleur, Ethel Julia, *An Analytical Study of Methods of Testing Art Appreciation*, Chicago: University of Chicago Press, 1925.

4) Bulley, M. H., "Art and the Child," *Burlington Magazine*, October, 1923, pp. 43–179.

5) Bulley, M. H., "An Enquiry as to Aesthetic Judgments of Children," *British Journal of Educational Psychology*, June, 1934, 4: 162–82.

6) Calkins, M. W., "An Experiment in Aesthetic Appreciation," *Psychology Review*, 1900.

7) Carroll, Herbert A. and Eurich, Alvin C., "Abstract Intelligence

and Art Appreciation," *Journal of Educational Psychology*, March, 1932, 23: 214–20.

8) Carroll, H. A., "Preliminary Report on a Study of the Relationship between Ability in Art and Certain Personality Traits," *School and Society*, August 27, 1932, 36: 285–88.

9) Carroll, H. A., "What Do the Meier-Seashore and the McAdory Art Tests Measure?," *Journal of Educational Research*, May, 1933, 26: 661–65.

10) Cattell, Judith: Gloscock, Josephine; Washburn, M. F., "Experiments on a Possible Test of Aesthetic Judgment of Pictures," *American Journal of Psychology*, 1918.

11) Christensen, Erwin O. and Karwoski, Theodore, "Test in Art Appreciation," Report I, University of North Dakota Department Bulletin, Vol. 10, No. 7, November, 1926.

12) Drought, R. Alice, "A Survey of Studies in Experimental Aesthetics," *Journal of Educational Research*, September, 1929, 20: 97–102.

13) Druley, Helen, *Traits of Children Gifted in Art*, Master's Thesis, Minnesota, 1935.

14) Eurich, A. C. and Carroll, H. A., *Educational Psychology*, Boston: D. C. Heath and Company, 1935, p. 435.

15) Eurich, A. C. and Carroll, H. A., "Group Differences in Art judgment," *School and Society*, August 8, 1931, 34: 204–6.

16) Garrells, Harriet E., *The Relationship between General Intelligence of Sixth-grade Pupils and Their Ability to Appreciate Art as Measured by the McAdory Test*, Master's Thesis, George Washington University, 1934.

17) Garrison, S. C., "Fine Arts," *Review of Educational Research*, December, 1934, 4: 498–500.

18) Gordon, K., "A Study of Aesthetic Judgments," *Journal of Experimental Psychology*, 6, 1923, pp. 36–43.

19) Jewell, Edward Alden, *Vollard-Comment on Career of Noted Dealer*, *New York Times*, Art Section, July 30, 1939.

20) Kinter, Madeline, *The Measurement of Artistic Abilities*, The Psychological Corporation, New York, 1933.

21) Knauber, Alma Jordon, "Testing for Art Ability," *Education*, December, 1934, pp. 219–33.

22) Longman, "Art Education in the University," *Parnassus*, October, 1936, 8: 21–25.

23) McAdory, Margaret, *The Construction and Validation of an Art Test*, Contributions to Education, No. 383, New York: Teachers College, Columbia University, 1929.

24) Meier, N. C., "Aesthetic Judgment as a Measure of Art Talent," *University of Iowa Studies*, Series 1, No. 14, 1926, pp. 3–30.

25) Miller, George T., *Reactions of Junior High School Students to the Architectonic of Pictures*, Master's Thesis, Pittsburgh, 1931.

26) Mursell, James L., "The Application of Psychology to the Arts," *Teachers College Record*, January, 1936, 37: 290–299.

27) Norton, Jennie A., *The Effect of One Year's Customary Classroom Art Work on the Aesthetic Judgment of Seventh-grade Pupils*, Master's Thesis, Pittsburgh, 1935.

28) Pan, S., "A Study in Esthetic Judgment: the Influence of Familiarity," *Psychological Abstracts*, No. 2134, April, 1936, 10: 233.

29) Peters, Charles C. and Van Voorhis, Walter R., *Statistical Procedures and Their Mathematical Bases*, State College, Pennsylvania: The Pennsylvania State College, 1935.

30) Pintner, Rudolf, *Aesthetic Appreciation of Pictures by Children*, Columbus, Ohio: Ohio State University, 1918.

31) Scott, Helen North, *An Evaluation of Two Types of Teaching for Appreciation of Art*, Master's Thesis, Indiana, 1932.

32) Speer, Robert K., *Measurement of Appreciation in Poetry, Prose, and Art, and Studies in Appreciation*, Contributions to Education, No. 362, New York: Teachers College, Columbia University, 1929, p. 77.

33) Stinson, Ruth Eby, *The Relation of Artistic Production and Aesthetic Sensitivity in Children*, Master's Thesis, Iowa, 1934.

34) Sutherland, Eula Ward, *A Determination of the Art Appreciation of Eighth-grade Girls in Story County, Iowa, as a Basis for Courses in Art Related to Homemaking*, Master's Thesis, Iowa State College, 1931.

35) Swartz, D. J., "Pint of Art," *Bulletin of High Points*, May, 1931, 13: 17–20.

36) Terry, Marion Baker, *Art Appreciation versus Technical Skills in Art in the Junior High School*, Master's Thesis, Southern California, 1933.

37) Tyler, R. W., "Review of the Construction and Validation of an Art Test by Margaret McAdory," *Educational Research Bulletin*, 9, 1930, pp. 501–2.

38) Voss, Mildred Dow, *A Study of Conditions Affecting the Functioning of the Art Appreciation Process at the Child Level*, Doctor's Thesis, Iowa, 1936.

39) Wall, Ernest A., *Aesthetic Sense and Education*, Doctor's Thesis, New York, 1931, 415 p. ms.

40) Wallis, Nell, *A Study of Tests Designed to Measure Art Capacities*, Master's Thesis, Florida State College for Women, Tallahassee, 1930.

41) Waymack, Eunice Hammer, *Child Reactions as a Basis for Teaching Picture Appreciation*, Master's Thesis, Cincinnatti, 1931.

42) Weldele, Mary K., *Art in the Life of the Public School Pupil*, Master's Thesis, Indiana State Teachers College, 1933.

43) Wentworth, Helen C., *Comparative Achievement of College Students as Measured by Three Art Appreciation Tests*, Master's Thesis, Minnesota, 1935.

44) White, R. and Johnson, B.: "Children's Choices in Modern Art," *Child Development*, I, 1930, pp. 347–349.

45) Whiteford, William G., "Empirical Study of Pupil-Ability in Public School Art Courses," *The Elementary School Journal*, xx, September and October, 1919, pp. 35 and 41.

WIMBERLY'S CRITICISMS OF KUHLMANN-ANDERSON TESTS

ROSE G. ANDERSON
Psychological Corporation

In the April 1946 issue of this JOURNAL Wimberly[1] reports the results of an exhaustive analysis of the Kuhlmann-Anderson Group Tests based upon the performance of seventy-seven ten-, eleven-, and twelve-year-old children and one hundred sixteen clinic referrals nine through thirteen years of age.

There is much interesting material in the article. The interpretations and conclusions, however, are based upon an erroneous understanding of the procedure used in arriving at mental age equivalents for the Kuhlmann-Anderson tests.[2] Wimberly presents results which he considers evidence of systematic errors in the mental-age equivalents and accounts for these presumed errors on the basis of the incorrect use of the regression of chronological age on test score in arriving at mental-age equivalents, whereas the approved method for arriving at such equivalents employs the regression of test score on chronological age.

Since the approved method was followed in arriving at the Kuhlmann-Anderson norms in question, it is unfortunate that the Kuhlmann-Anderson *Manual* is not completely clear on this score. This lack of clarity has resulted in considerable misplaced emphasis in Wimberly's study.

Wimberly arrives at his conclusions on the basis of two statements in the Kuhlmann-Anderson *Manual* which, taken out of context, are admittedly ambiguous. Neither statement should have been misunderstood by Wimberly, however, in view of other references in the *Manual* to Kuhlmann-Anderson norms.

The basis for arriving at mental-age equivalents has been indicated in the *Manual* in the paragraph directly preceding one of the statements upon which Wimberly bases his interpretation[2] (p. 2). In explanation of the graphic representation of the mental age ranges for the tests the statement is made. "The left ends of the heavy horizontal lines in the chart indicate the median ages of the children who pass one trial and no more in each of the tests and the right ends of these lines indicate the median ages of the children up to eighteen years who pass all

trials." That is, mental-age equivalents reading vertically down the graph under age six are based upon the median performance of unselected six-year olds for each of the designated tests and similarly for each successive age represented in the graph.

The method of arriving at mental age equivalents is further indicated by the following statement[2] (p. 8): "We propose to measure mental development from the age of five to mental maturity. For this purpose that test is most valid which shows this development best, by having the highest rate of increase through successive years . . . the age norm table gives a rough indication of the discriminative capacity of each test . . . the thirty-nine tests . . . were selected . . . because they gave the largest and most consistent increases in score for successive age levels."

Throughout the *Manual* the importance of representative school populations and of unselected age groups received repeated emphasis. Elsewhere[3] the present writer listed six precautions essential to assuring the highest possible validity of tests, two of which are "a sufficiently large number of likely tests so that only those having the highest discriminative capacity need be retained" and "retention of those which discriminate best between unselected successive age groups."

In the context of the above statements it is puzzling that Wimberly concludes that mental ages have been arrived at on the basis of the 'averaged' performance of children of a range of ages arbitrarily agreed upon by the authors. He makes this interpretation from the statement[2] (p. 2) under 'Standardization': "the average age was determined at which one trial and no more, two trials and no more, etc. were passed." This statement is historical. It describes the preliminary exhaustive analysis each test was exposed to in the selection of the final tests (from over a hundred) for the battery. In every case a test was given to a wide range of ages and for each age the per cents of children passing 1 to N trials was computed. Those tests were considered eligible for inclusion which showed the largest consistent increases in the per cents passing any given number of trials from age to age and also consistent increases within each given age in per cents passing increasing numbers of trials.

At the time of the original standardization (early twenties) the tests were given to a much broader range of ages than was later found necessary to determine the upper range of applicability; that is, the maximum mental age the test measured. This is the gist of the second statement from which Wimberly concludes that the authors arbitrarily excluded certain ages in arriving at the 'averaged' performance which determined the norms.

Since the statements in question were misunderstood by Wimberly they admittedly need clarification. The 'mental-set' resulting from his interpretation of them has led to some devious reasoning.

Much of what Wimberly has gone to considerable effort to demonstrate is inherent in the structure of the scale and the principles underlying its construction. The objective results are favorable to both.

On the basis of Kuhlmann's statement that "as many or as few tests may be used" in an individual examination "as time and judgment dictate" Wimberly uses ranges of twenty tests as the basis for his analysis of the performance of the three age groups—ten, eleven, and twelve. He overlooks the implication that an individual examination would be given by a clinically qualified person whose 'judgment' would dictate that tests be selected on the basis of their capacity to measure the individual concerned. In his choice of tests Wimberly used ten tests in addition to the ten tests indicated by the *Manual* according to the grade placements of the children. Among the added tests for the ten-year-olds he has added two tests with maximum mental ages of nine-six and ten-three, respectively; for the eleven-year-olds he has added the above two tests and three tests with maximum mental ages of eleven-ten, eleven-seven, and eleven-ten, respectively. For the twelve-year-olds he has added two tests with maximum mental ages of eleven-seven and eleven-ten. From the above the following predictions can be made: an increase of approximately ten per cent of maximum scores for above average ten-year-olds; an increase of ten per cent maximum scores for average and below average eleven-year-olds, and an increase up to fifty per cent of maximum scores for superior eleven-year-olds; an increase of ten per cent of maximum scores for above average twelve-year-olds. It is

not surprising, therefore, that Wimberly finds an increase of 9.74 per cent in maximum scores for the added ten tests over the recommended ten tests. Nor is it surprising that he finds an increase of only 1.69 per cent of zero scores, since in the remaining added tests the mental-age equivalents for one trial passed in the two most difficult tests are 10-0 and 10-9, respectively, for the ten- and eleven-year-olds, 12-0 and 10-11, respectively, for the twelve-year-olds. A small increase in zero scores would be expected for the below-average children in these groups.

What is surprising is the significance Wimberly attaches to his findings, i.e., that the "absolute percentages are of little value in determining whether the added tests lowered the difficulty of the battery." This doubt persists in spite of the above mental age data; in spite of his own finding that the mean IQ for the twenty tests was lower than the mean IQ for the ten tests recommended for the grade placement of the child; in spite of his own finding that the mean IQ for the ten easy tests was 89.65, and for the ten hard tests was 103.42 in comparison with a mean IQ of 99.62 for the ten tests recommended for grade placement; and in spite of his own conclusions that the larger gains in IQ from the easier to the harder tests were associated with higher intelligence.

The proportion of maximum scores and the higher mean IQ on the ten most difficult tests both indicate that the ability of a certain proportion of above average children had not been exhausted by the booklet corresponding to their grade placement. This contingency has been considered by the authors[2] (p. 12): "If, on the other hand, the graph lies near the upper limits of most of the tests with perhaps two or more maximum scores, the indication is that the next higher booklet is more suitable to his level of ability . . . the child should be given the remaining tests in the other booklet and be scored on the basis of his performance in the tests comprising the more appropriate booklet."

Arguing that IQ's gained on more difficult tests are too high is exactly analogous to basing a child's Stanford-Binet IQ on performance through the average adult level in spite of the fact that he is capable of passing tests at the superior adult levels.

Wimberly's comparison of Kuhlmann-Anderson and Stanford-

Binet IQ's for clinic referrals are more relevant with respect to the clinic's use of the Kuhlmann-Anderson test than they are with respect to relationships between the two scales. He states: "the operative clinical procedure in most cases dictated the administration of the Kuhlmann-Anderson battery recommended for the grade placement of the child."

It is the 'expected average mental level' of the group or individual which should guide the selection of the Kuhlmann-Anderson booklet to be used. Using the test booklet corresponding to the child's grade placement and stopping there is analogous to using the same range of Stanford-Binet tests for all fourth-graders or all ten-year-olds. It violates both the letter and the spirit of the instructions[2] (pp. 1, 12)[3] (p. 200). In the use of the test for individual examination by clinicians, the latter have relevant case history data and clinical judgment to guide them in adapting the tests used to the child's expected intelligence level. That such a basis for judgment was available is implied in Wimberly's statement: "the majority of children seen by the Bureau represents an instance of academic difficulty which had resulted in school placements below mental age expectancy." Yet an arbitrary procedure was used which inevitably resulted in inadequate measurement for a proportion of the children involved.

On the basis of Wimberly's figures the per cent of zero scores for his clinical group is 2.14 times as large as for his unselected ten-, eleven-, and twelve-year-olds. the per cent of maximum scores, 1.54 times as large. Because of zero or maximum scores Kuhlmann-Anderson mental ages were computed on the basis of only five test scores (instead of on ten) in two cases, on six test scores in two cases, on seven test scores in three cases, on eight test scores in fourteen cases. Further, in eight of these cases there are combinations of zero and maximum scores which provide evidence of erratic performance or inadequate test administration. From the above data, it is obvious that the tests were not well chosen, and that further testing was necessary in order to arrive at representative mental ages. It was also advisable in the twenty-one additional cases whose mental ages were based upon nine test scores.

The results secured for the clinical group are thus seen to be below the minimal results obtainable by the recommended

group examination procedure. No effort has been made to exhaust the resources of the tests by following the recommended individual examination procedure. Consequently, conclusions with respect to comparisons between Kuhlmann-Anderson and Stanford-Binet IQ's have limited claim to validity. Under the necessity of having his interpretations conform to the 'rationale' of his study, Wimberly attributes discrepancies found to errors in the Kuhlmann-Anderson tests.

No comment is made here with respect to Wimberly's criticisms of various statements in earlier articles by Kuhlmann which have no bearing on the 1942 Kuhlmann-Anderson norms for which the present writer was solely responsible.

Acknowledgment is made herewith of the desirability of clarifying the statements in the Kuhlmann-Anderson *Manual* which have been shown to be subject to misunderstanding.

REFERENCES

1) Wimberly, Stan E. "A Systematic Error in Kuhlmann-Anderson Mental Ages," *J. Educ. Psychol.*, 1946: 4, 193–218.

2) Kuhlmann-Anderson Tests: *Instruction Manual* (Fifth Edition). Minneapolis: Educational Test Bureau, 1942.

3) Anderson, Rose G. "Fifth Revision of Kuhlmann-Anderson Test," *J. Appl. Psychol.*, 1940: 2, 198–206.

INFLUENCE OF DIFFICULTY OF PERCEPTUAL-MOTOR TASK UPON ACCELERATION OF CURVES OF LEARNING

WILLIAM C. F. KRUEGER

Wayne University, Detroit, Michigan

This report presents a further study of the general problem concerning the rate of progress as related to the difficulty of the assignment.[1] The experiment was designed to observe the relationship between varying degrees of difficulty of perceptual-motor tasks and the corresponding acceleration of learning curves obtained from the practice of the various tasks.

In this series of experiments the learners were required to toss rings (Mason jar rings), one at a time, in such a manner that the rubber ring should fall upon and remain hanging on a nail in the wall, four feet above the floor. The four groups of subjects used were undergraduate college students. Group I (twenty subjects) pitched the rings from a distance of two feet. The distance was determined by measuring the intervening space between the toe of the subject's shoe nearest to the wall and the wall proper. The subjects were required to pitch the rings and not merely to lean forward to hang the rings on the nail. Group II (twenty subjects) tossed the rings for a distance of three feet, while Group III and Group IV (each also containing twenty subjects) threw rings for a distance of six and nine feet, respectively. The various distances were to represent the various degrees of difficulty.

One trial consisted of tossing ten rings, one ring at a time. The score for any one trial was the number of rings that remained hooked. One preliminary practice session was given to each subject to familiarize him with the conditions of the experiment. Thereafter the trials for the experiment proper began. Ten successive trials were completed in any one practice session, and two practice periods were permitted in any one day. There was always an interval of at least four hours between two practice periods. The subjects continued to work on successive days until each subject had completed one hundred trials. Group I,

[1] Krueger, Wm. C. F., "Rate of Progress as Related to Difficulty of Assignment," *Jour. Edu. Psychol.*, 1946, Vol. 37, pp 247–249.

however, practiced only for twelve trials because all subjects had made perfect scores on successive trials by the time the twelfth trial was completed. No subject participated in more than one learning condition. An attempt was made to equals the groups as much as was feasible.

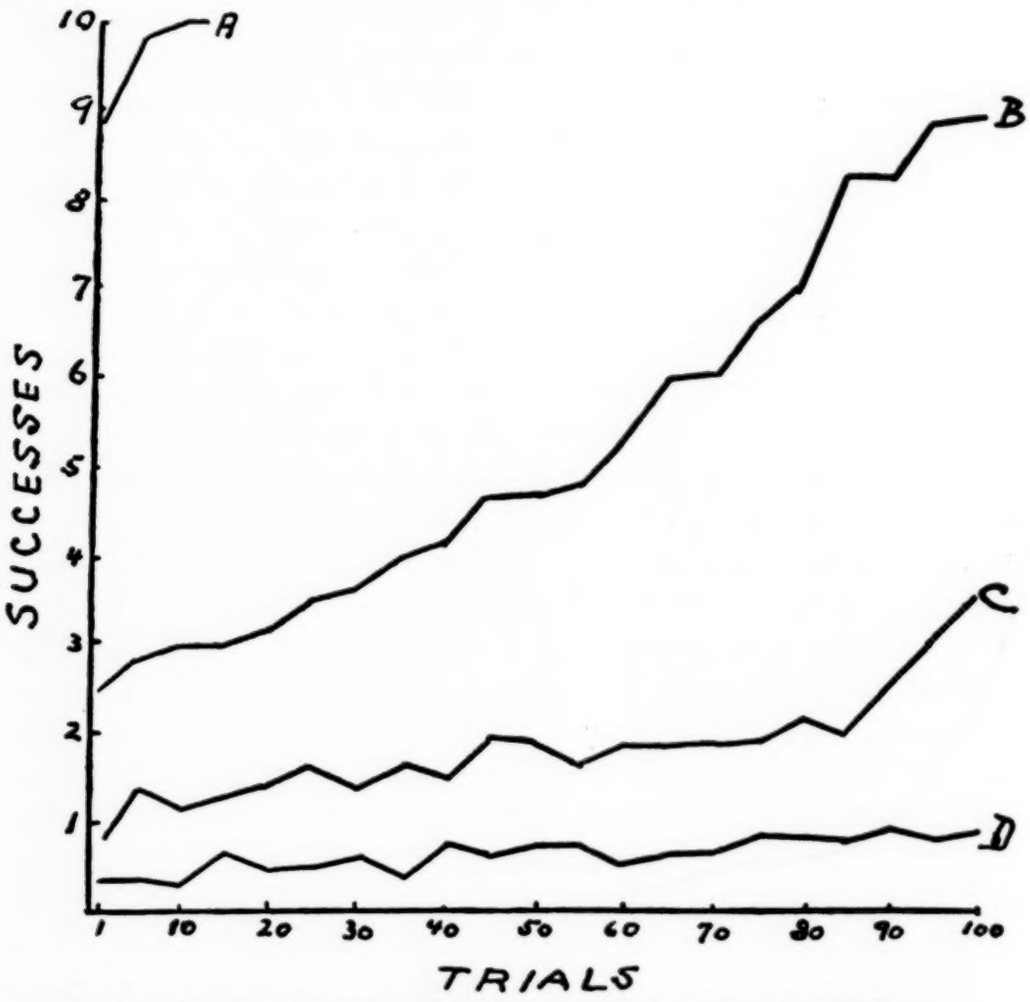

FIG. 1. Curves of Learning Derived from Tasks of Varying Difficulty. Key: (A) Distance of two feet. (B) Distance of three feet. (C) Distance of six feet. (D) Distance of nine feet.

The averages for the various trials and the various condition of learning are presented in Fig. I. By inspection it may be noted that as the distance for the tossing of the rings was increased, the number of successes decreased. When the assigned task was very simple, as in the two-foot distance, the subjects learned quickly and soon achieved perfection. The rate of learning was negatively accelerated. When the throwing distance was increased to three feet we found that the early trend yielded an

almost straight line with possibly a trace of positive acceleration. Toward the end of that curve there was a slight shift toward a negative trend of acceleration. When the distance was increased to six feet, the rate of accleration was definitely more positive than before, with the upward sweep of the curve beginning noticeably toward the end of the practice periods. The longest distance, nine feet, showed almost no gain for the entire practice period, although the average scores during the last trials were somewhat higher than those obtained during the initial stages of the learning period.

The averages for the first, fifth, tenth, fifteenth and every successive fifth practice periods were used to plot the respective curves.[1]

CONCLUSIONS

Four curves of learning were obtained. Although they show different rates of acceleration, they may show four different phases of a general curve of learning. We may say that we found the presence of this trend: When the task assigned is very simple, we may expect a negatively accelerated curve of learning. When the requirement is rather difficult for the learner, the consequent progress will follow a pattern of positive acceleration. As the task shifts from one that is relatively easy for the learning to one that is relatively difficult, corresponding rate of learning will shift from negative to positive acceleration.

[1] Because of the restrictions at present placed on special matter, the detailed table of average learning scores for the various conditions is omitted. The author will be glad to supply mimeographed copies to those desiring them.

BOOK REVIEWS

ALLEN L. EDWARDS, *Statistical Analysis for Students in Psychology and Education*. New York: Rinehart and Co., 1946, pp. 360.

The assumption is made that a textbook in statistics is primarily a teaching aid, and that consequently the major justification for a new text in this area—already provided with a number of such teaching aids—is that it will prove to be a better means to the ends that should be accomplished by the teaching of statistics to students of psychology and education. In advertising this text, the publisher has emphasized its 'modern approach' and its fitness for use with 'the non-mathematically-trained student.' That the 'non-mathematically-trained student' is often with us is evident to teachers of elementary statistics. Any text that enables us to help these students—as well as other students—to achieve the appropriate concepts, skills, understandings, abilities, and attitudes has much to recommend it. Whether or not the principle of beginning one's teaching at the point that the student's present development warrants is strictly a modern principle may be questioned; certainly, though, it is a common-sense one. With respect to mathematical training, the author assumes that students are not familiar with arithmetic and algebraic operations. In an early chapter he explains and illustrates rules for these operations; further, he refers the student to these rules at points later in the book where algebraic operations are being used to derive formulas. This assumption that students are not familiar with simple mathematical procedures necessarily emphasizes the elementary character of the text, particularly of the early sections that deal with descriptive statistics.

In preparing a text, the author necessarily defines the field by the act of including certain topics and excluding others. He points out in the preface the omission of certain topics. These are partial and multiple correlation, the reliability and validity of tests, and the "usual extended treatment of centiles and scaling techniques." One also notes the absence of the familiar discussion of various graphical methods of presenting data. Topics that are included are: measures of central tendency and variability (Chapter 3), product-moment coefficient of correlation (Chapter 5), other measures of association (Chapter 6), prob-

ability and frequency distributions (Chapter 7), sampling distributions (Chapter 8), the *t* test (Chapter 9), analysis of variance (Chapters 10 and 11), Chi-square (Chapter 12), prediction (Chapter 13), and research and experimentation (Chapter 14).

The introductory chapter suggests ways in which the student may use the text and introduces major terms and functions of statistical method. Chapter 2 is the mathematical introduction, and Chapter 4 discusses coding and grouping techniques that may be used to simplify calculations. According to the preface, partial and multiple correlation are omitted because the author has: " . . . searched without much success to find many applications of these statistical techniques in the literature. That they have their uses as well as their limitations is fully recognized, but I feel that they may be developed more profitably . . . in advanced courses, where time spent on these topics will not be at the expense of statistical techniques which are more commonly in use."

As the above list of topics suggests, there is emphasis on small sample theory. This is defended in the preface: "In discussing tests of significance, I have emphasized small sample theory since, whether the traditional attitude approves or not, more and more research as published in the psychological and educational journals is being evaluated by small sample techniques."

The criterion of use as a basis for the inclusion and the exclusion of topics is clearly evident. The assumption that present practice—apparently determined by some method of quantification, as the terms 'commonly in use' and 'more and more' imply —provides a desirable guide to the selection of topics is implicit.

Another important aspect of a text in elementary statistics—and 'text' implies that it may be the only volume on statistics that many of the students will ever read—is the concept of the kinds of problems that statistical methods permit us to attack. The author has drawn on a number of actual research studies as illustrations of what can be done with the techniques discussed. 'His realism' is to be admired; even a number of his hypothetical examples are so designed that they give results consistent with actual studies. Four major functions of statistical method are discussed in the opening chapter: making precise descriptions, planning and designing experiments, determining the degree of confidence that can be placed in certain generalizations made

from observations, and making predictions. In the final chapter, the author suggests kinds of questions that have been investigated and discusses such aspects of design as the concept of the null hypothesis, kinds of sampling error, sample size, and methods of selecting control groups.

The emphasis in the text is on kinds of problems that can be attacked by comparing measures for groups that are known to differ in certain respects. Tests of significance constitute the focus of this volume. The treatment of *t*, analysis of variance, and Chi-square is the best exposition in the book. The author follows the practice of referring the student to other volumes for the mathematical development; consequently, the treatment has an 'applied' rather than a 'theoretical' character. Many critical comments and explanations that supplement the main body of the discussion appear in footnotes; a number of these warrant greater emphasis than their appearance in footnote form gives them. Analysis of variance has been treated critically; for example, epsilon-square, which Peters and Van Voorhis suggest be used to replace analysis of variance, is included in the discussion of matched-group analyses.

The early sections of the book are less satisfactory. Chapter 3, for example, introduces in a total of twenty-two pages the following concepts: range, arithmetic mean, average deviation, variance and standard deviation, the normal distribution curve, median, semi-interquartile range, centiles, standard scores, and 'other measures of central tendency and variability.' A very large proportion of these twenty-two pages is devoted to the operations involved in calculating the statistics that are presented; ample statements about their meaning and use are absent. At least seven of nine paragraphs devoted to the median deal primarily with its calculation; one of the remaining two paragraphs considers its appropriate use. And this paragraph, it might be noted, emphasizes only that the median may be appropriately used "to supplement the description provided by the mean" if the distribution is markedly skewed.

The treatment of product-moment correlation also emphasizes calculation procedures. The danger of interpreting correlation as a cause-effect relationship is pointed out, but the purposes for which r is to be used are given a scanty treatment. The rule-of-thumb procedure for evaluating r as very high, high, moderate,

low, and very low is prominent in the discussion of interpretation of correlation. Concern about the 'non-mathematically-trained student' apparently has led to detailed exposition of how these statistics are computed at the expense of indicating what they mean. If understanding of the meaning and use, as well as the calculation, of commonly-used measures of central tendency, variability, and association is an important outcome of the elementary course in statistics, the present text must be judged to be an inadequate means to this end.

CHESTER HARRIS
Department of Education
University of Chicago

ROBERT S. GILCHRIST and Faculty of the University School. *How Children Develop.* (A revision of *Child Development Study.*) Columbus: The Ohio State University, 1946, pp. 79.

Developmental data are classified under seven age levels: three to six, six to nine, nine to eleven, eleven to thirteen, thirteen to fourteen, fourteen to sixteen, sixteen to eighteen. Statements designed to give a picture, presumably a cross section, of development at each of these age levels are grouped under the following headings:

I. Maintaining Personal Health and Promoting Healthful Living
II. Achieving and Maintaining a Sense of Security
III. Developing and Maintaining a Sense of Achievement
IV. Developing and Maintaining Ever-widening and Deepening Interests and Appreciations

Likewise, a series of uniform subheads, but different for each main head, is employed for all age levels. The subheads for III are:

A) Gaining a Sense of Personal Adequacy and Orientation
B) Understanding and Gaining Control over the Environment
C) Developing Sensitivity to the Opposite Sex

Statements about development varying in number from three or four to about ten appropriate to each subhead are provided, as for example, III B age six to nine: "7. Children of these ages can solve simple everyday problems of living, such as writing and

mailing a letter, and going to the store and making purchases"; and III B age nine to eleven: "9. These children can be expected, with training, to take care of their own rooms and help with other chores about the house. They are capable . . . of assuming responsibility for personal needs in dress and grooming."

The statements appear to be well chosen. Within limits of the facets of growth about which statements are made, one may gain an appreciation of the developmental expectation at each age level, subject, of course, to the validity of the statements. Individual differences are not stressed. The reader can, by following through a given major head and its subheads for the various age levels, gain a fair picture of the course of development for a given problem, such as "Developing and Maintaining a Sense of Achievement."

Abundant references are found at the end of each age period. Presumably, it is intended that these supply data on the dynamics and mechanisms of development. The text is enlivened with a number of illustrative cartoons.

J. B. STROUD
State University of Iowa

ASAHEL D. WOODRUFF. *The Psychology of Teaching.* New York: Longman's, Green and Company, 1946, pp. 180.

This book represents a significant attempt to tie up educational psychology with practical problems which teachers are daily facing and which prospective teachers need to consider. Woodruff is convinced, as are many others, that "educational psychology the country over is making a very meagre contribution to the work of the average schoolteacher." He further believes that educational psychology texts are suited primarily for advanced graduates majoring in educational psychology. He contends that courses in educational psychology should be primarily of a laboratory nature if the work is to be functional. The book is a relatively short one since the author believes prospective teachers need to learn a few facts, with their main efforts devoted to learning how to apply these facts to practical problems.

Although chapters on growth and development and on the problems of childhood and adolescence are omitted, most of the chapter headings are similar to those in more conventional

educational psychology tests. The emphasis in most subjects tends to differ significantly, however, from that in conventional texts. There is considerable attention given to the part played by motivation and motivational health in the promotion of effective learning. The various types of learning are emphasized and the discussion of the steps in effective problem-solving as the learner strives toward his goal is quite different from the summary of learning experiments frequently given in educational psychology textbooks.

The teacher who is wedded to the conventional text would tend to shy away from this book because it does not contain the specific results of classic and near-classic experiments so dear to the educational psychologist. Woodruff, by implication, questions whether such listing of experimental results is the best way to get teaching influenced by educational psychology.

The educational psychology teacher who agrees with Woodruff that there is relatively little 'carry-over' value of educational psychology from the conventional text will probably wish that the author had departed farther from the conventional pattern and actually started with the problems which the beginning teacher faces and had shown how psychology could contribute to their solution. In some chapters, this is done to a degree.

While in most chapters the author assumes that the learner shall share in the making of decisions in the classroom, in other chapters such as "The Nature of Mental Activity" there is the implied assumption that practically all decisions shall be made by the teacher. It is also assumed in the chapter on "Evaluating Status and Progress" that all evaluating shall be done by the teacher. No learner-self-evaluation seems to be a part of the author's philosophy. Thus, there seems to be some inconsistency with advocacy of much learner participation at some points and the opposite emphasis in other sections of the book.

The reviewer feels the analysis of many more specific classroom situations and problems, particularly in the first two-thirds of the book, would have materially strengthened the possibility of students later applying the theoretical principles so well analyzed. The author perhaps feels these can be supplied by the teacher.

Some of the most significant areas discussed are: motivational health; learner goals, learner values, basic steps needed in functional problem-solving, and learner maladjustment and its causes.

The book has definite merit as an effort, which is at least partially successful, to make educational psychology realistic and functional.

RAY H. SIMPSON
University of Alabama

HELEN M. ROBINSON. *Why Pupils Fail in Reading.* Chicago: University of Chicago Press, 1946, pp. 257.

The clinical worker or teacher who is dealing with reading disability cases is confronted with a large array of anomalies which have been emphasized as causes of disability by one writer or another. Adequate remedial treatment is possible only when there is tangible evidence that an anomaly is responsible for part or all of the reading difficulty exhibited by a case. The first section of this book includes a discussion of significant contributions dealing with causes of reading retardation. Areas covered include visual maladjustments, neurological factors, auditory and speech difficulties, physical deficiencies, intelligence, emotional and personality problems, and environmental and social characteristics of poor readers. Some of these survey chapters are very well done—for instance the discussion of emotional and personality problems. In general, the technique followed is to cite and coördinate the conclusions arrived at by writers rather than a critical analysis of experimental results.

In the second major part of the book the author reports a study designed to identify and to evaluate the causal factors in a group of thirty severely retarded readers with IQ's ranging from 85 to 137. A social worker, a psychiatrist, a pediatrist, a neurologist, three ophthalmologists, a speech specialist, an otolaryngologist, an endocrinologist, a reading specialist and a psychologist coöperated in examining the cases, in evaluating the anomalies founded, and in identifying possible causes of reading retardation. An intensive remedial program was carried out with twenty-two of the cases to secure evidence of the potency of each of the possible causes. Among the findings and the conclusions are the following: (1) The more seriously retarded pupils presented the greater number of anomalies. This suggests that difficulty in reading may be a part of the general deviation from the normal pattern. (2) Evaluation by the group of specialists indicated that some of the anomalies discovered for

a particular child were coincidental and had no direct relationship to his reading deficiency. (3) A number of factors that appeared to be possible causes of reading failure, did not prove experimentally to be so. That is, the mere presence of an anomaly does not mean that it is contributing to reading failure. (4) The most frequent causes of reading disability were social, visual, and emotional difficulties. Less important were inappropriate school methods, neurological factors, speech and functional auditory difficulties. Least important of all were endocrine disturbances, general physical difficulties and insufficient auditory acuity. The author is well aware that these frequencies do not apply to reading disability populations in general. Her group was highly selected. For the most part, they consisted of problem cases brought in voluntarily by parents seeking aid. This kind of selection probably explains to some degree the high frequency of cases who had emotional maladjustment and social problems listed as causes of reading disability. Furthermore, since little information was available concerning teaching methods to which the children had been exposed, 'school methods' as a causal factor must have received less emphasis than it deserved.

The careful reader will suspect that in some instances the assigning of a factor as a probable cause is not entirely valid. For instance, Case 4 needed a small visual correction which was furnished. He was then given concentrated help with reading, using special methods adapted to his needs. Satisfactory gains were made. In the final evaluation, visual difficulty was assigned as one of the probable causes of the reading disability. With a case of this kind, it seems to the reviewer that it is questionable whether the minor visual difficulty was a contributing cause.

The author takes the fundamentally sound viewpoint that reading disability is usually caused by a pattern of several interacting factors rather than by a single anomaly. Since most clinicians cannot call in a group of specialists to aid in diagnosis, it is suggested that there should be a thorough search for symptoms of various causal factors. Then, if a symptom appears, the appropriate specialist may be consulted. Also a sound emphasis is shown by the suggestion that "a direct, vigorous reading program must follow correction of causal factors."

Adequate remedial techniques in reading consist of no more or less than good teaching of reading with an emphasis upon indi-

vidualized instruction. This implies that the instruction must be adjusted to the needs (determined by adequate diagnosis) of the child. Apparently, if motivation is adequate, any one of several methods (Gates, Fernald, Monroe, etc.) will yield good results. The unique contribution of this book is that the author provides valuable suggestions, based upon experiment, for making a more valid diagnosis of causes of reading disability in the clinical situation. The volume is noteworthy both as a research contribution and as a source of practical suggestions concerning the causes of reading retardation.

MILES A. TINKER
University of Minnesota

PUBLICATIONS RECEIVED

GRACE ARTHUR. *Tutoring as Therapy.* New York: The Commonwealth Fund, 1946, pp. 125. $1.50.

ROGER G. BARKER, BEATRICE A. WRIGHT, AND MOLLIE R. GONICK. *Adjustment to Physical Handicap and Illness: A Survey of the Social Psychology of Physique and Disability.* New York: Social Science Research Council, 1946, pp. 372. $2.00.

FRANZISKA BAUMGARTEN. *Die Psychologie der Menchenbehandlung im Betriebe.* Zurich: Rascher Verlag, 1946, pp. 304.

JOSEPH CAHN AND OTHERS. *The Story of a Discussion Program: Veterans and Their Neighbors Get Together on Public Issues.* New York: New York Adult Education Council, 1946, pp. 94. (paper) $1.00.

EDGAR DALE. *Audio-Visual Methods in Teaching.* New York: (386 Fourth Ave.) The Dryden Press, 1946, pp. 546. $4.25.

MARY FORD. *The Application of the Rorschach Test to Young Children.* Minneapolis: University of Minnesota Press, 1946, pp. 114. $2.00.

WILLIAM H. FOX. *Spelling Proficiency in Township Schools in Indiana.* Bloomington, Ind.: Indiana University, 1947, pp. 32. (paper) $0.50.

HENRY E. GARRETT. *Statistics in Psychology and Education.* Third Edition. New York: Longmans, Green and Co., 1947, pp. 465. $4.00.

ARTHUR I. GATES AND ROSE E. KUSHNER. *Learning to Use Hearing Aids: A Study of Factors Influencing the Decision of Children to Wear Hearing Aids.* New York: Bureau of Publications, Teachers College, Columbia University, 1946, pp. 77. (paper)

KENNETH B. HAAS AND HARRY Q. PACKER. *The Preparation and Use of Visual Aids.* New York: Prentice-Hall, Inc., 1946, pp. 224. $4.00.

ARTHUR T. JERSILD AND ASSOCIATES. *Child Development and the Curriculum.* New York: Bureau of Publications, Teachers College, Columbia University, 1946, pp. 274. $2.75.

CLYDE KLUCKHOHN AND DOROTHEA LEIGHTON. *The Navaho.* Cambridge, Mass.: Harvard University Press, 1946, pp. 258. $4.50.

GEORGE LAWTON. *Aging Successfully.* New York: Columbia University Press, 1946, pp. 266. $2.75.

KATHARINE M. MAURER. *Intellectual Status at Maturity as a Criterion for Selecting Items in Preschool Tests.* Minneapolis: The University of Minnesota Press, 1946, pp. 166. $2.50.

MARIAN J. RADKE. *The Relation of Parental Authority to Children's Behavior and Attitudes.* Minneapolis: University of Minnesota Press, 1946, pp. 123. $2.00.

FRANCIS P. ROBINSON. *Effective Study.* New York: Harper and Brothers, 1946, pp. 262.

G. MILTON SMITH. *A Simplified Guide to Statistics for Psychology and Education.* Revised and Enlarged. New York: Rinehart and Co., Inc., 1946, pp. 109. (paper)

HENRY LESTER SMITH, KATHLEEN DUGDALE, BEULAH FARIS STEELE, AND ROBERT STEWART MCELHINNEY. *One Hundred Fifty Years of Grammar Textbooks.* Bloomington, Indiana: Indiana University, 1946, pp. 199. (paper) $0.50.

EDITH M. STERN AND HOWARD W. HOPKIRK. *The Housemother's Guide.* New York: The Commonwealth Fund, 1946, pp. 91. (paper) $0.50.

ALLYS DWYER VERGARA. *A Critical Study of a Group of College Women's Responses to Poetry.* New York: Bureau of Publications, Teachers College, Columbia University, 1946, pp. 159. $2.35.

HELEN LELAND WITMER, Editor. *Psychiatric Interviews with Children.* New York: The Commonwealth Fund, 1946, pp. 443. $4.50.

WERNER WOLFE. *The Personality of the Preschool Child.* New York: Grune and Stratton, 1946, pp. 341. $5.00.